ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

FACULTAD DE INFORMÁTICA Y ELECTRÓNICA

CARRERA TELECOMUNICACIONES

"ESTUDIO DE LOS PROTOCOLOS BGP EVPN/VXLAN COMO MECANISMO DE INTERCONEXIÓN DE DATA CENTERS"

Trabajo de Integración Curricular
Tipo: Proyecto de Investigación

Presentado para optar el grado académico de:
INGENIERO EN ELECTRÓNICA, TELECOMUNICACIONES Y REDES

AUTOR:

DUVAL ANDRÉS MENA PAREDES

Riobamba – Ecuador

2022

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

FACULTAD DE INFORMÁTICA Y ELECTRÓNICA

CARRERA TELECOMUNICACIONES

"ESTUDIO DE LOS PROTOCOLOS BGP EVPN/VXLAN COMO MECANISMO DE INTERCONEXIÓN DE DATA CENTERS"

Trabajo de Integración Curricular
Tipo: Proyecto de Investigación

Presentado para optar el grado académico de:
INGENIERO EN ELECTRÓNICA, TELECOMUNICACIONES Y REDES

AUTOR: DUVAL ANDRÉS MENA PAREDES

DIRECTOR: Ing. ALBERTO LEOPOLDO ARELLANO AUCANCELA MSc.

Riobamba – Ecuador

2022

Yo, Duval Andrés Mena Paredes, declaro que el presente trabajo de titulación es de mi autoría y los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autor asumo la responsabilidad legal y académica de los contenidos de este trabajo de titulación; el patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

Riobamba, 08 de marzo de 2022

Duval Andrés Mena Paredes
160065694-4

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

FACULTAD DE INFORMÁTICA Y ELECTRÓNICA

CARRERA TELECOMUNICACIONES

El Tribunal del Trabajo de Integración Curricular certifica que: El trabajo de Integración Curricular; tipo: Proyecto de Investigación, **ESTUDIO DE LOS PROTOCOLOS BGP EVPN/VXLAN COMO MECANISMO DE INTERCONEXIÓN DE DATA CENTERS**, realizado por el señor: **DUVAL ANDRÉS MENA PAREDES**, ha sido minuciosamente revisado por los Miembros del Trabajo de Integración Curricular, el mismo que cumple con los requisitos científicos, técnicos, legales, en tal virtud el Tribunal Autoriza su presentación.

	FIRMA	FECHA
ING. PEDRO SEVERO INFANTE MOREIRA Dr.C. **PRESIDENTE DEL TRIBUNAL**	 Firmado electrónicamente por: PEDRO SEVERO INFANTE MOREIRA	08/03/2022
ING. ALBERTO LEOPOLDO ARELLANO AUCANCELA MSc. **DIRECTOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR**	 Firmado electrónicamente por: ALBERTO LEOPOLDO ARELLANO AUCANCELA	08/03/2022
ING. PAÚL PATRICIO ROMERO RIERA MSc. **MIEMBRO DEL TRIBUNAL**	 Firmado electrónicamente por: PAUL PATRICIO ROMERO RIERA	08/03/2022

DEDICATORIA

A mis padres quienes sin descanso lo han dado todo por verme convertido en un profesional, a mi hermana Grace quien ha sido un gran apoyo a lo largo de mi carrera estudiantil, a mi compañera de vida Bella, quien con su paciencia, amor y alientos me ha dado ánimo para no rendirme y a mi gran amor mi hija Sara, porque todo sacrificio busca brindarle un mejor futuro.

Duval

AGRADECIMIENTO

Agradezco a Dios que me ha brindado sabiduría y perseverancia para salir victorioso frente a las adversidades que se presentaron a lo largo del camino, a mis padres que han sido el pilar fundamental, porque su sacrificio no fue en vano y son quienes merecen todo el crédito de este triunfo.

Agradezco a mi esposa, la cual ha tenido la paciencia y el tino para animarme en los momentos difíciles, es quien me ha dado el regalo más grande que es mi hija Sara, son mi felicidad y el motor que me impulsa para seguir adelante.

Agradezco al Ing. Alberto Arellano quien me ha brindado su sabiduría para culminar con éxito el trabajo de titulación, su guía y profesionalismo fueron intachables.

Por último, agradezco a la Dra. Mirian Jurado por todo el aprecio y los consejos que me ha brindado, es una persona muy importante para mí y para mi familia.

Duval

TABLA DE CONTENIDO

CAPÍTULO II

CAPÍTULO III

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

ÍNDICE DE GRÁFICOS

ÍNDICE DE ABREVIATURAS

AFI: Identificadores de la familia de direcciones

BGP: Protocolo de pasarela fronteriza

DC: Centro de datos

DCI: Interconexión de centros de datos

ECMP: Enrutamiento de rutas múltiples de igual costo

EVI: Instancia de EVPN

EVPN: Red privada virtual Ethernet

FTP: Protocolo de transferencia de archivos

IRB: Enrutamiento y puenteo integrados

MTU: Unidad máxima de transferencia

RD: Distintivo de ruta

RT: Objetivos de ruta

SAFI: Identificadores de la familia de direcciones subsiguientes

STP: Protocolo de árbol de expansión

VLAN: Red de área local virtual

VNI: Identificador de red VXLAN

VRF: Enrutamiento y reenvío virtual

VTEP: Punto final del túnel VXLAN

VxLAN: Red de área local extensible virtual

WAN: Red de área amplia

RESUMEN

El propósito de este proyecto fue analizar el rendimiento de los protocolos EVPN/VxLAN como una solución para interconectar centros de datos, partiendo por el establecimiento de las ventajas y desventajas que presenta la arquitectura Clos frente a la arquitectura tradicional para centros de datos, fue necesario realizar un estudio de cada uno de los protocolos que intervienen en la interconexión con el objetivo de conocer sus propiedades, de esta forma se diseñó dos topologías de red comparativas con diferentes tecnologías de virtualización para interconectar tres centros de datos por medio de un proveedor de servicios configurado con MPLS, luego apoyándose de un servidor FTP, se realizó un conjunto de 50 pruebas para paquetes de 1MB, 50MB y 100MB con la finalidad de obtener parámetros de rendimiento en función del ancho de banda, latencia, jitter y pérdida de paquetes, posteriormente se realizó un análisis comparativo del cual se pudo notar que al usar tramas gigantes existe una mejora notable en cuanto al rendimiento del ancho de banda lo que beneficia a disminuir los problemas ocasionados por jitter y pérdida de paquetes; sin embargo, en cuanto a los valores de latencia en EVPN/VxLAN se evidencio un efecto negativo por parte de la sobrecarga que genera el encabezado de VxLAN en relación a los valores que se obtienen en las VLANs con la MTU estándar, debido a que los equipos que forman parte de la topología hacen uso de un mayor porcentaje de recursos para el procesamiento de las tramas causando que los valores de latencia aumenten, sin embargo al realizar la categorización en base a los valores recomendados por los estándares internacionales, se estableció que los valores de latencia se encuentran dentro de la categoría “Excelente”, por lo que no representan un factor negativo en la interconexión de centros de datos.

Palabras clave: <CENTRO DE DATOS> <ARQUITECTURA CLOS> <RED PRIVADA VIRTUAL ETHERNET (EVPN)> <RED DE ÁREA LOCAL EXTENSIBLE VIRTUAL (VXLAN)> <PARÁMETROS DE RENDIMIENTO>.

SUMMARY

The purpose of this project was to analyze the performance of EVPN/VxLAN protocols as a solution to interconnect data centers, starting with the establishment of the advantages and disadvantages of Clos architecture compared to traditional architecture for data. For the investigation it was necessary to carry out a study of each of the protocols that intervene in the interconnection with the aim of knowing its properties, in this way it designed two comparative network topologies with different virtualization technologies for interconnect three data centers through a service provider configured with MPLS, next relying on an FTP server, a set of 50 tests was carried out for packets of 1MB, 50MB and 100MB in order to obtain performance parameters based on the bandwidth, latency, jitter and packet loss; however, regarding the latency values in EVPN/VxLAN a negative effect was evidenced by the overhead generated by the VxLAN header in relation to the values obtained in the VLANs with the standard MTU, because the teams that are part of the topology make use of a greater percentage of resources for the processing of frames causing latency values to increase. However, when carry out the categorization based on the values recommended by international standards, it was established that the latency values are within the "Excellent" category, therefore they do not represent a negative factor in the interconnection of data centers

Keywords: <DATA CENTER> <CLOS ARCHITECTURE> < ETHERNET VIRTUAL PRIVATE NETWORK (EVPN)>, <VIRTUAL EXTENSIBLE LOCAL AREA NETWORK (VXLAN)> <PERFORMANCE PARAMETERS>.

Firmado electrónicamente por:
WILSON GONZALO
ROJAS YUMISACA

MSc. Wilson G. Rojas
NOMBRE Y FIRMA PROFESOR

INTRODUCCIÓN

Antecedentes

Los líderes de infraestructura y operaciones (I&O) enfrentan hoy un desafío abrumador. La TI que conocen desde hace décadas está teniendo un cambio radical. La función principal de TI será permitir negocios más ágiles, ingresar a nuevos mercados de forma más rápida, brindar servicios más cerca del cliente y posicionar cargas de trabajo específicas en función de los impactos comerciales, regulatorios y geopolíticos. La función del centro de datos tradicional se está relegando a la de un área de almacenamiento heredada, dedicada a servicios muy específicos que no se pueden admitir en otro lugar, o que respalda los sistemas que son más eficientes económicamente en las instalaciones (Cappucio, 2018).

A medida que los servicios de interconexión, los proveedores de nube, el Internet de las cosas (IoT), los servicios de borde y las ofertas de SaaS continúen proliferando, la razón para permanecer en una topología de centro de datos tradicional tendrá ventajas limitadas. Este no es un cambio de la noche a la mañana, sino un cambio evolutivo en la forma de pensar en cómo se brinda los servicios a clientes y al negocio (Cappucio, 2018).

"La transformación digital, la aparición de nuevos servicios y el crecimiento nunca visto de los flujos de datos han cambiado para siempre el entorno digital. Cada vez más empresas recurren a soluciones de 'Data Center Interconnect' (DCI) para proporcionar a sus centros de datos conexiones estables, fiables y de banda ultra ancha" (Datta, 2020).

En este nuevo mundo, factores como el rendimiento y la latencia son críticos, pero también lo son una operación y mantenimiento inteligentes y sencillos, reforzando el ámbito de la seguridad. Como resultado, la tecnología de interconexión de centros de datos (DCI) se está volviendo cada vez más importante, debido a su capacidad para aumentar el ancho de banda del centro de datos, reducir la latencia y eliminar la pérdida de paquetes, lo que lleva a una mejor experiencia general (Computerworld, 2020).

DCI es una solución que permite la intercomunicación entre las VM de varios DC. Utilizando tecnologías como VXLAN junto a BGP EVPN, DCI transmite de forma segura y confiable paquetes de comunicación a través de redes portadoras. Con DCI, la intercomunicación entre las VM de varios DC es posible independientemente de si estas VM residen en la misma VLAN (Huawei, 2021).

Formulación del problema

¿El uso de los protocolos BGP EVPN/VXLAN como mecanismo DCI mejorará el rendimiento de conexión frente a soluciones tradicionales?

Sistematización del problema

- ¿Cuáles son los mecanismos existentes para la Interconexión de Data Centers?
- ¿Qué relación existe entre los protocolos BGP EVPN/VXLAN como mecanismo de interconexión de Data Centers?
- ¿Cuáles son las ventajas y desventajas de utilizar los protocolos BGP EVPN/VXLAN como mecanismo DCI?
- ¿Cuáles son los principales parámetros de rendimiento que deben ser utilizados en mecanismos DCI?

Justificación teórica

Tradicionalmente, los centros de datos usaban muchos enlaces de capa 2 que abarcaban racks completos, filas, jaulas y pisos, hasta donde alcanzaba la vista. Estos grandes dominios L2 no eran ideales para un centro de datos debido a la lentitud de la convergencia, las transmisiones innecesarias y la dificultad de administración. Para optimizar la red del centro de datos, necesitábamos reducir el uso y la dependencia de los protocolos de capa 2, como Spanning Tree. Sin embargo, el desafío es el hecho de que los centros de datos necesitan que la capa 2 se extienda de un estante a otro, de una fila a otra, a veces de un centro de datos a otro, no solo para los requisitos de la aplicación, sino también para la tolerancia a fallas y la movilidad de la carga de trabajo. Han surgido numerosas tecnologías para combatir esta limitación, como TRILL, FabricPath y VXLAN. De estos tres, es Virtual Extensible LAN (VXLAN) que ha tenido una rápida adopción en los centros de datos modernos (Varnum, 2018).

VXLAN define un esquema de tunelización para superponer redes de Capa 2 sobre redes de Capa 3. Permite el reenvío óptimo de tramas Ethernet con soporte para múltiples rutas de tráfico unidifusión y multidifusión con el uso de encapsulación UDP / IP para tunelización, y se utiliza principalmente para la conectividad de sitios dentro del centro de datos. Por otro lado, una característica única de EVPN es que el aprendizaje de direcciones MAC entre los conmutadores

ocurre en el plano de control utilizando MP-BGP, este método difiere de las soluciones VPN de capa 2 existentes, como VPLS, que aprenden inundando unidifusión desconocida en el plano de datos. Este método de aprendizaje MAC basado en el plano de control es el habilitador clave de las muchas funciones útiles que ofrece EVPN (Juniper, 2021).

La tecnología EVPN proporciona mecanismos para la interconexión del centro de datos (DCI) de próxima generación al agregar procedimientos de plano de control extendido para intercambiar la información de la capa 2 (dirección MAC) y la capa 3 (dirección IP) entre los enrutadores fronterizos del centro de datos (DCBR) participantes. Estas características ayudan a abordar algunos de los desafíos de DCI, como la movilidad perfecta de la VM y el enrutamiento IP óptimo (Juniper, 2021).

Los resultados de este estudio permitirán relacionarse con el ámbito de las TI, las dificultades que se presentan en los negocios relacionados con centros de datos, e ir aprendiendo de nuevos métodos de conexión que se desarrollan para combatir las necesidades de las comunicaciones y el mundo global interconectado.

Justificación aplicativa

En el presente proyecto se analizarán los protocolos BGP EVPN/VXLAN que permiten la interconexión de Centros de datos que resulta como un método de solución basado en la virtualización de redes como consecuencia del avance tecnológico y la adopción masiva de máquinas virtuales, lo que ha provocado limitaciones en cuanto al número permisible de VLANs en un centro de datos, que se traduce como un factor limitante en función de la escalabilidad; de esta forma con la ayuda de una plataforma de simulación de topologías de red llamada GNS3, se procederá a diseñar una topología que interconecte tres centros de datos por medio de un proveedor de servicios MPLS apoyándose de las plataformas de enrutamiento de las marcas Arista Networks y Cisco, por consiguiente con la ayuda de un servidor configurado con el protocolo FTP, se pretende estresar a la topología por medio de descargas con la finalidad de obtener parámetros de rendimiento, permitiendo así conocer los beneficios de aplicar tecnologías nuevas enfocadas a la ampliación de los horizontes en centros de datos.

Objetivos

Objetivo General

Estudiar los protocolos BGP EVPN/VXLAN como mecanismo de interconexión de Data Centers.

Objetivos Específicos

- Realizar una comparación entre arquitecturas de Centros de datos tradicionales versus arquitecturas de Centros de Datos actuales, protocolos, ventajas y desventajas.
- Analizar las funciones de los protocolos BGP, EVPN y VXLAN, su relación y su adaptación en entornos DCI.
- Definir los principales mecanismos y protocolos utilizados en la Interconexión de Data centers.
- Diseñar una topología para DCI bajo los protocolos EVPN, VXLAN mediante una plataforma de simulación de red.

CAPÍTULO I

1. MARCO TEÓRICO REFERENCIAL

1.1. Centro de datos

La necesidad de un mundo interconectado y el rápido crecimiento de Internet en la década de 1990 impulsó una demanda de infraestructura de red que operaba continuamente y permitía a las organizaciones mantener una presencia constante en línea. Tanto proveedores de servicios de Internet como las empresas de nivel empresarial promovieron la construcción de redes e infraestructura de centros de datos que pudieran brindar servicios, soporte y capacidad de almacenamiento (Gyarmathy, 2019).

Un centro de datos descrito de la forma más simple y concisa, hace referencia a una amplia gama de equipos e instalaciones físicas que las empresas utilizan para alojar sus aplicaciones y datos críticos. El diseño de un centro de datos está enfocado en una red de recursos informáticos y de almacenamiento que permiten la entrega de aplicaciones, brindar soporte y resguardar información primordial de las organizaciones. Los componentes que forman parte del diseño de un centro de datos incluyen firewalls, conmutadores, enrutadores, sistemas de almacenamiento, servidores y controladores de entrega de aplicaciones (Cisco, 2019). Formando así la columna vertebral de Internet y son indispensables para las telecomunicaciones y los servicios digitales modernos (Gyarmathy, 2019).

1.1.1. Importancia y características de un centro de datos

La ola creciente de la tecnología en diferentes ámbitos hace que las personas consuman y produzcan datos de forma masiva y a un ritmo nunca visto: Internet, motores de búsqueda, aplicaciones para móviles, teléfonos inteligentes, etc. Por ello se considera que la tecnología se ha vuelto una pieza esencial en el diario vivir de personas, grupos y empresas. Pero la realidad es que todos los artilugios tecnológicos actuales, y que están en constante desarrollo, todas las actividades empresariales enfocadas a la interconexión, dependen del almacenamiento, la distribución en red y el procesamiento de datos digitales; todas estas aplicaciones tienen como destino ser atendidas y ejecutadas por un centro de datos o con su mediación. Absolutamente, los centros de datos son un pilar fundamental y los héroes anónimos de Internet, por lo que se han convertido en un sector de gran importancia para las organizaciones que ejecutan aplicaciones críticas (Glinkowski, 2013, p.7).

Por ende, para la mayoría de organizaciones, los Data Centers son el activo y componente más valioso, representando incluso la propia razón de su existencia y crecimiento. En ellos se busca mantener y unificar todos los servicios disponibles en un lugar determinado, así como toda su información y datos corporativos de relevancia. Resultando así de gran importancia para que las empresas estén al día en torno a las tecnologías que las mantienen no sólo en funcionamiento, sino respaldadas (KIO Networks, s.f, p.2).

Sullivan (2016), considera que un centro de datos exitoso debe cumplir con las siguientes caracteristicas:

- **Eficiencia en el espacio y planificación de la capacidad:** Los centros de datos están diseñados para lograr una densidad óptima, utilizando tecnologías de virtualización para ofrecer la mayor potencia informática, de almacenamiento y de red con el mínimo espacio. Aumentando servidores, reduciendo bastidores y con ello enfocarse en mejorar el hardware existente gracias a la disminución de gastos por bastidor.
- **Escalabilidad/ Flexibilidad:** La flexibilidad en la infraestructura de los centros de datos permite una mayor capacidad de soporte de actualizaciones y de mantener un rápido crecimiento en cuanto a rendimiento. El objetivo primordial es permitir que nuevas tecnologías y sistemas puedan ser adaptadas al centro de datos con el menor tiempo de interrupciones.
- **Eficiencia de Enfriamiento:** Con un sistema de refrigeración adecuado se marca una gran diferencia con respecto a eficiencia energética en un centro de datos, son varios los métodos que se pueden aplicar, sin embargo, el modelo pasillo caliente/ pasillo frío reduce el consumo de energía para la refrigeración y aumenta el tiempo de vida útil de los dispositivos que forma parte del centro de datos.
- **Documentación, planificación y procedimientos:** Procedimientos bien documentados permiten tener un control riguroso y confiable sobre las operaciones del centro de datos, cumpliendo así las políticas establecidas por los organismos reguladores; la formación, la comunicación, la colaboración y la mejora continua son normas recomendadas para una buena gestión dentro del centro de datos.
- **Seguridad:** Gracias a la correcta documentación de los métodos y procedimientos, se logra garantizar la seguridad de las instalaciones de misión crítica, ya que impiden el acceso no autorizado. Los centros de datos más seguros cuentan con una construcción balística, protección perimetral, diseño adaptable con armarios seguros o entornos enjaulados y otras medidas de control de acceso de varios niveles.

1.2. Arquitecturas de centros de datos – Topología fija basada en arboles

1.2.1. Basic Tree: Access-aggregation-core

La arquitectura de centro de datos de tres niveles está compuesta por conmutadores de acceso conectados a los servidores, conmutadores de agregación que permiten la conexión a los conmutadores de acceso y los conmutadores del núcleo del centro de datos que proporcionan el enrutamiento hacia y desde la red central de la empresa. El diseño de tres niveles se basa en un diseño jerárquico, de tal forma que se pueden añadir nuevos pares de conmutadores de agregación/distribución sin necesidad de modificar los pares de agregación existentes (Enterasys Networks, 2013, p.8).

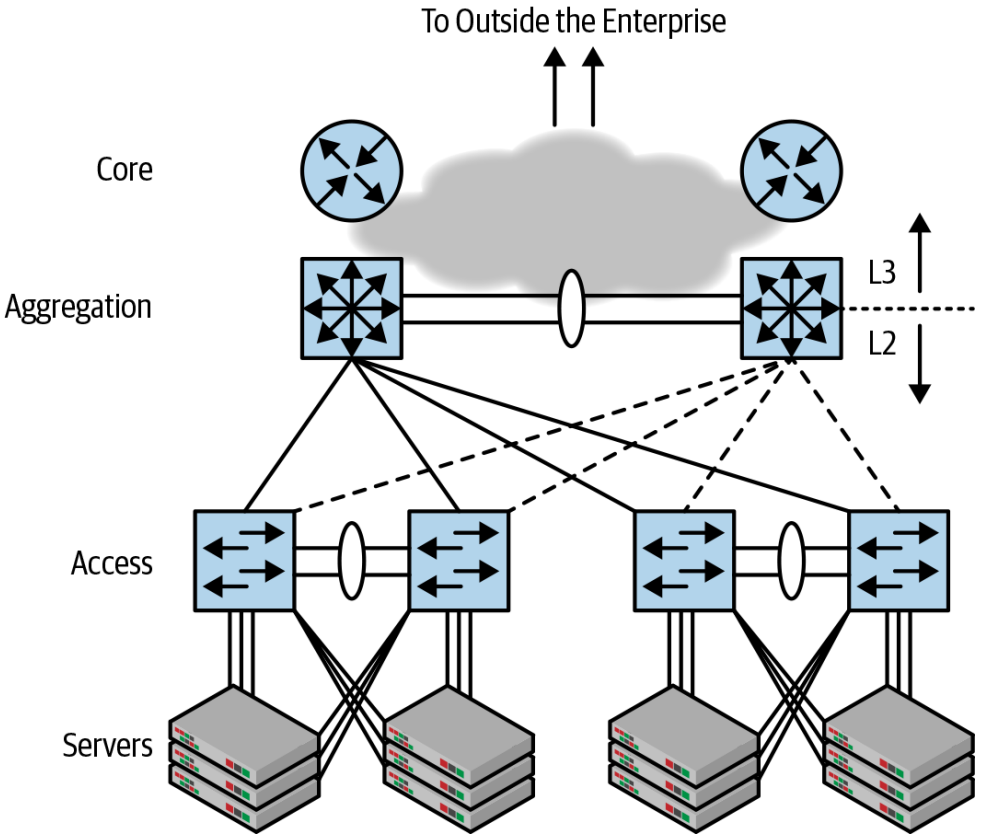

Figura 1-1: Arquitectura de Centro de Datos de tres niveles

Fuente: (Dutt, 2019)

Este diseño fue pensado y promovido por Cisco Networks en la década de 1990. El diseño se basó en patrones de tráfico actuales, las limitaciones en el diseño de equipos de última generación de ese período y la necesidad de seguridad (Marschke et al., 2011, pp.34-35).

La Figura 1-1, muestra los niveles de acceso, agregación y núcleo que forman parte de la topología basic tree; Al-shawi (2015), describe de la siguiente manera las funciones de cada una de las capas:

- **Capa central (Core):** Proporciona un transporte óptimo entre sitios y un enrutamiento de alto rendimiento. Debido a la criticidad de la capa central, los principios de diseño del núcleo deben proporcionar un nivel adecuado de resiliencia que ofrezca la capacidad de recuperarse rápidamente y sin problemas después de cualquier evento de fallo de la red con el bloque central.
- **Capa de agregación (Aggregation):** Proporciona conectividad basada en políticas y control de límites entre las capas de acceso y núcleo.
- **Capa de acceso (Access):** Proporciona el acceso de los grupos de trabajo y usuarios a la red.

1.2.1.1. Spanning Tree Protocol (STP) como una alternativa de solución dentro de la arquitectura basic tree.

En las redes cliente-servidor del siglo pasado, IP era uno de los cuantos protocolos de red para la interconexión, tecnologías como IPX o VINES eran utilizadas para la comunicación, el problema radicaba en que estos protocolos no funcionaban entre sí, cada uno poseía una arquitectura propia, no obstante, tenían una característica en particular, utilizaban el puente (bridging) para su funcionamiento. Por lo tanto, la arquitectura access-aggregation-core permitió a los ingenieros de red construir una red común para todos estos protocolos de red heterogéneos en lugar de construir una red diferente para cada tipo específico de protocolo (Dutt, 2019, pp.5-9).

Debido al escalado en las redes, el enrutamiento IP se volvió un tema complicado ya que implicaba una gran cantidad de configuración explicita, ya que los dos extremos de una interfaz deben configurarse para estar en la misma subred para que el enrutamiento funcione; además que, en cuanto a consumo de CPU el puente (bridging) ocupaba menos recursos que el enrutamiento IP. Por lo tanto, bridging prometía un diseño de red único para todos los protocolos de capa superior junto con una conmutación de paquetes más rápida y una configuración mínima. La realidad es que la creación de puentes tiene varias limitaciones como consecuencia tanto del modelo de aprendizaje como del Protocolo de árbol de expansión (STP) (Dutt, 2019, pp.5-9).

De este modo, si al enviar un paquete a un destino que no está en la red o uno que nunca se estableció, los puentes nunca sabrán dónde está este destino. Por lo tanto, incluso en una topología de triángulo simple, incluso con verificación de reenvío automático, el paquete rodeará el triángulo formando así un bucle infinito. Una notoria diferencia entre el encabezado IP del encabezado MAC, es que este último no contiene un campo de tiempo de vida (TTL) para evitar que un paquete se convierta en un bucle infinito. Incluso un solo paquete de transmisión en una red pequeña con un bucle puede terminar usando todo el ancho de banda disponible y dejando inservible a toda la red. A esta catástrofe se la denomina tormenta de transmisión (Dutt, 2019, p.9).

Por ello para evitar tal catástrofe la arquitectura access-aggregation-core hace uso del protocolo de árbol de expansión (STP).

Spanning Tree Protocol, asegura el funcionamiento de la red sin tormentas a costa de la eficiencia de los enlaces. El STP se recupera de los fallos de los enlaces mediante una serie de temporizadores que controlan el estado de los enlaces y los nodos. Una vez que estos temporizadores expiran, la red realiza el proceso de convergencia superando así los fallos por bucles infinitos. El tiempo de convergencia puede provocar el retraso de las comunicaciones durante 10 segundos o más. Aunque el RSTP puede menorar el tiempo de convergencia, la cantidad de tiempo que se necesita para reconocer un fallo y recuperarse no es aceptable en las redes actuales (Marschke et al., 2011, pp.35-37).

STP y los sucesores como RSTP no podían cumplir con las exigencias cambiantes de las redes, debido a sus múltiples desventajas y limitaciones, era necesario tomar medidas que pudieran dejar en el pasado los problemas asociados con STP, tecnologías de superposición como Fabric Path y TRILL pasaron a primer plano, introduciendo redes de Capa 2 enrutadas con una encapsulación de superposición MAC-in-MAC. Esto se convirtió en una superposición de MAC en IP con la invención de VXLAN. Si bien las redes de capa 2 evolucionaron más allá de las topologías sin bucles con STP, las funciones de puerta de enlace de primer salto para la capa 3 también se volvieron más sofisticadas. Las puertas de enlace centralizadas tradicionales alojadas en las capas de distribución o agregación han pasado a implementaciones de puertas de enlace distribuidas. Esto ha permitido escalar y eliminar puntos de estrangulamiento (Krattiger et al., 2017).

1.2.1.2. Uso de VLANs en la arquitectura access-agg-core.

El costo de utilizar STP para evitar el riesgo de bucles en la red es demasiado alto, hay que recordar que los conmutadores de acceso están conectados a dos conmutadores de agregación y mantener la mitad de enlaces suspendidos es una solución retrograda. Por ello, Cisco desarrollo una tecnología denominada árbol de expansión por VLAN (PVST), permitiendo funcionar ambos enlaces entre los conmutadores de acceso y agregación; para que esto se logre, se creó un árbol de expansión por cada VLAN. Al colocar las VLAN pares en uno de los conmutadores de agregación y las VLAN impares en el conmutador adyacente, el protocolo permitía el uso de ambos enlaces: un enlace por el grupo par de VLAN y el otro por el grupo impar, logrando así aumentar el ancho de banda en la red (Dutt, 2019, p.11).

1.2.2. Clos: Spine-Leaf

Clos, es una topología basada en árboles que ha tenido una amplia acogida gracias a la integración de conmutadores de red de alta velocidad. La arquitectura Clos tiene una regla bastante simple y eficiente: los conmutadores del nivel "x" deben conectarse sólo a los conmutadores de los niveles "x-1" y "x+1", pero nunca a otros conmutadores del mismo nivel. Son varias las ventajas que presenta la arquitectura Clos, entre ellas: un buen porcentaje de resiliencia, un alto grado de redundancia, adecuado reparto de la carga de tráfico y tolerancia a fallos; gracias a la redundancia de las rutas entre pares de conmutadores se logra una utilización eficiente de los recursos de la red. Además, permite un enorme ancho de banda de bisección, de forma que este ancho de banda de bisección es el mismo en todos los niveles, por lo tanto, no existe una sobresuscripción, lo que resulta eficiente para determinadas aplicaciones. También debido a su forma es rápida para la resolución de problemas de tráfico y evita la carga añadida de tener un núcleo y una capa de agregación separados que fomenta la arquitectura tradicional de tres niveles (Kapadia et al., 2014, pp.6-8).

La Figura 2-1, muestra el diseño de la arquitectura CLOS de dos niveles, en la que una serie de conmutadores de agregación (denominados espinas) dan cabida a una serie de conmutadores de acceso (denominados hojas). En el diseño se puede contemplar que 32 conmutadores de columna están conectados a 256 conmutadores de hoja de 48 puertos, obteniendo así un tejido CLOS de centro de datos capaz de dar servicio a 12.288 dispositivos de borde (Kapadia et al., 2014, pp.8).

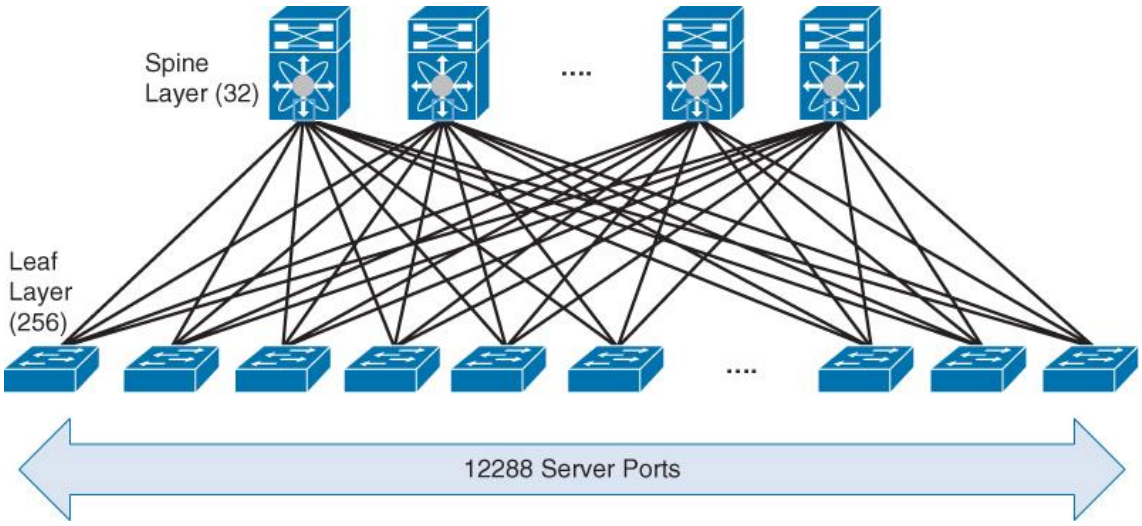

Figura 2-1: Arquitectura Clos (Spine-Leaf)

Fuente: (Kapadia et al., 2014, p.8)

Gracias a su diseño, Clos permite una red de alta capacidad, dado que hay más de dos rutas entre dos servidores cualesquiera. El agregado de espinas permite aumentar el ancho de banda disponible para las hojas. Pero no es aconsejable agregar más enlaces entre una hoja y una espina

con el fin de ganar más ancho de banda. El único propósito de las espinas es dar interconexión a las diferentes hojas. Por ello las espinas no brindan ningún otro servicio y tampoco se comunican con los puntos finales del centro de datos. Por lo tanto, las espinas son diferentes de los conmutadores de agregación de la arquitectura access-aggregation-core descrita en el apartado 1.2.1., aunque estructuralmente ocupan la misma posición. En otras palabras, en una topología de Clos, toda la funcionalidad se empuja a los bordes de la red, las hojas y los servidores mismos (Dutt, 2019, p.20).

En la red access-aggregation-core solo está permitido trabajar con dos conmutadores de agregación lo cual es una notable limitación. La arquitectura Clos puede admitir más de dos espinas (conmutadores de agregación), esto se logra debido a que el Protocolo de árbol de expansión (STP) no se utiliza como protocolo de control de interconexión de conmutadores. Además, el puente (bridging), se utiliza principalmente solo en los bordes, es decir, dentro de un solo bastidor. Para lograr la conexión entre racks, se utiliza una solución moderna y con mayor eficiencia, esta se denomina red de área local extensible virtual (VXLAN), esta se describe con mayor detalle en el apartado 1.4 (Dutt, 2019, p.22).

La industria de las redes de centros de datos, probó diversos mecanismos antes de decidirse por una solución definitiva con el enrutamiento IP, la idea era conservar el puenteo como un mecanismo que permitiera trabajar sin el uso del protocolo STP; tecnologías como TRILL, FrabicPath de Cisco, Shortest Path Bridging (SPB), se implementaron en algunas redes; sin embargo, estas soluciones eran inmaduras. Además, el puente había dejado de ser útil como método de conectividad principal porque las redes IP eran prácticamente el único protocolo de capa superior en uso (Dutt, 2019, p.22).

De este modo, la arquitectura Clos permite el enrutamiento con más de dos espinas apoyándose en el protocolo de enrutamiento de múltiples rutas de igual costo (ECMP). Esto permite que un paquete se reenvíe a lo largo de cualquiera de las rutas de igual costo disponibles. De esta forma una hoja puede alcanzar cualquier otra hoja usando cualquier espina, y el costo para alcanzar cualquier hoja es el mismo (Dutt, 2019, p.22).

Cada hoja posee una o más subredes IP, y actúa como puerta de enlace por defecto para estas subredes. Las subredes pueden estar en el mismo o en diferentes VRFs. Los protocolos de enrutamiento de capa 3 que soportan ECMP se utilizan entre las hojas y las espinas, principalmente BGP, con algún despliegue de OSPF o IS-IS (Gai, 2020).

1.2.2.1. Arquitectura Clos de tres niveles: Super-Spines

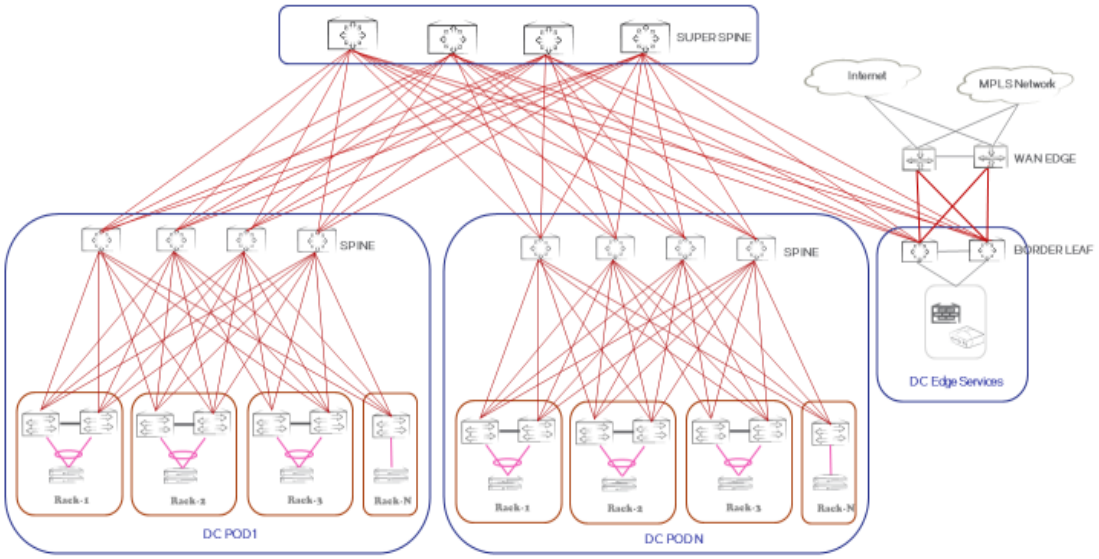

Figura 3-1: Arquitectura Clos de tres niveles con super espinas.

Fuente: (Extreme Networks, 2018, p.16)

La arquitectura Clos de tres niveles se construye tomando dos puertos de los conmutadores de espina para luego conectarlos a otra capa de conmutadores (la capa superior), como se muestra en la Figura 3-1. Este diseño está pensado para centros de datos de tamaño mediano a muy grande, siendo así que empresas como Microsoft, Amazon y muchos otros operadores han adoptado esta arquitectura de red. Este diseño se denomina comúnmente modelo de pod o clúster (Dutt, 2019, p.29).

Esta topología añade un nuevo nivel a la red, conocido como super espina. Las super espinas funcionan de forma similar a las espinas: Reenvío en el plano de control y en el plano de datos de BGP entre los PODs y desde los PODs a destinos fuera del tejido a través de las hojas de borde. No hay puntos finales conectados a las super espinas. La Figura 3-1 muestra cuatro switches de super espina que conectan los conmutadores de espina a través de múltiples PODs de centros de datos (Extreme Networks, 2018, p.16).

La conexión entre las espinas y las super espinas siguen los principios de Clos:

- Cada columna vertebral (espina) se conecta a todas las super espinas de la red.
- Tanto las espinas como las super espinas no están interconectadas entre sí (Extreme Networks, 2018).

1.2.3. Comparación entre la arquitectura Access-aggregation-core vs Spine-leaf: ventajas y desventajas.

Tabla 1-1: Ventajas y desventajas de las arquitecturas Access-aggregation-core vs Spine-leaf.

Arquitectura Basic Tree: Access-aggregation-core	
Ventajas	**Desventajas**
- Uso de mecanismos de puenteo (bridging)/puentes de autoaprendizaje, que permitían a varias tecnologías trabajar en la arquitectura Access-agg-core. - Uso de spanning tree protocol (STP), para evitar las tormentas de transmisión, es decir, una topología de árbol sin bucles. - Aumento del ancho de banda a través del árbol de expansión por VLAN. - Admite múltiples protocolos de capa superior. - Se considera una arquitectura rápida, barata, más simple de administrar y adecuada para el tráfico norte-sur. - Permite la calidad de servicio (QoS). - Esta arquitectura permite la escalabilidad.	- Si se inyecta un paquete a un destino que no está en la red, los puentes nunca sabrán dónde está este destino, por lo tanto, incluso en una topología simple, se generará una tormenta de transmisión. - Un mayor ancho de banda de este-oeste se logra incorporando conmutadores de agregación; sin embargo, STP no permite trabajar con más de 2 conmutadores de agregación; por lo tanto, un ancho de banda limitado significa que la red sufre de congestión. - Aprender un millón de direcciones MAC con el método flood and learn de los puentes de autoaprendizaje se considera como anti escalable, ya que los tiempos de espera son inviables en cualquier topología de red. - Las escalas actuales de las nubes (clouds), hicieron que 4096 VLANs se consideren como un número insignificante para satisfacer la demanda dentro de los centros de datos. - La repetida latencia de enrutamiento asociada al diseño de tres niveles no es aceptable para las comunicaciones de servidor a servidor de baja latencia. - La seguridad debe integrarse en todas las capas del diseño, no sólo en los bordes.

	- La baja utilización de los enlaces asociada al STP y a las rutas alternativas redundantes es antieconómica. - Los costes de escalabilidad del diseño son muy elevados.
Arquitectura Clos: Spine-Leaf	
Ventajas	**Desventajas**
- Gracias a Charles Clos, hoy en día se puede construir una red de conmutación muy grande y económica, utilizando simples conmutadores de factor de forma fijo. - La eliminación del protocolo STP, permite que la arquitectura spine-leaf pueda usar más de dos conmutadores espinas esto gracias a que se utiliza una solución de virtualización moderna denominada VXLAN y adopta el protocolo ECMP para el enrutamiento. - Menor consumo de energía usando múltiples enlaces de 10 GbE entre la columna y la hoja en lugar de enlaces de 40 GbE. Además, al usar enlaces de 10 GbE, se puede obtener un máximo de 5832 servidores y con enlaces de 40 GbE solo 512. También ayuda en el equilibrio de carga y al perder un enlace de 10 GbE produce una reducción menor en el ancho de banda que perder uno de 40 GbE. - Con más de dos espinas, la pérdida de un solo enlace o un nodo espina no es catastrófica. Los grandes proveedores de escala web y algunos otros utilizan hasta 16 o 32 espinas. Es decir, la pérdida de un solo nodo o enlace de espina da como resultado una reducción de solo 1/16 del ancho de banda total. - Preparada para el futuro para una red de mayor rendimiento a través de la selección de	- Los conmutadores de forma fijo crean la necesidad de gestionar muchos cables, esto supone un gran problema en función del escalado en la red, ya que muchas veces el cableado es una causa común de fallos en la red, y da como resultado la reducción del rendimiento en el centro de datos; es necesario adoptar tecnologías de verificación de cables para mitigar este problema. - Debido a la presencia de múltiples conmutadores de factor de forma fijo, la configuración de la red ya no es una tarea sencilla, de hecho, es todo lo contrario a tener una arquitectura agradable; Por lo tanto, la automatización de la red se convierte en algo indispensable. - El número de hosts que se pueden admitir puede ser limitado debido a que el recuento de puertos de la columna vertebral restringe el número de conexiones de los conmutadores de hoja. - Puede producirse una sobresuscripción de las conexiones hoja-espina debido al número limitado de conexiones en la espina disponibles para los conmutadores hoja (normalmente de 4 a 6). Por lo general, no se considera aceptable una relación de sobresuscripción de 5:1 entre la hoja y la

conmutadores de mayor cantidad de puertos en la columna vertebral. - Escalado sencillo de nuevos racks añadiendo nuevos conmutadores Leaf en la parte superior del rack. - Puesto que el diseño espina-hoja pertenece a la arquitectura Clos, se puede escalar aún más agregando un nivel superior y dando como resultado un modelo espina-hoja-super espina; obteniendo así cabida para un mayor número de usuarios y mayor ancho de banda en la red.	espina, pero esto depende en gran medida de la cantidad de tráfico en su entorno particular.

Fuente: (Dutt, 2019**;** Shamsee et al., 2015**;** Nelson, 2017)

Realizado por: Mena Duval, 2021.

1.3. Enrutamiento IP

El enrutamiento de paquetes IP, es el conjunto de tareas necesarias que requiere una red para trasladar un paquete IP de un enrutador a otro hasta llegar a su destino.

1.3.1. BGP

El protocolo de puerta de enlace fronteriza (BGP), está definido en el RFC 1654, se caracteriza por ser un protocolo de vector de ruta y es considerado como un protocolo de puerta de enlace exterior (EGP), BGP permite escalabilidad, flexibilidad y estabilidad de red. Los objetivos que han sido considerados para BGP desde su creación principalmente son: la inter organización para conectividad IPv4 en redes públicas, Internet, o redes privadas dedicadas (Jain y Edgeworth, 2016, pp.1-2).

Donohue y Stewart (2010, pp.69-70), explican las características que destacan en BGP:

- Los enrutadores que ejecutan BGP se denominan altavoces BGP.
- Los sistemas autónomos (AS), son un concepto propio de BGP, este a su vez se describe como un grupo de redes que se rige por una administración en común. La IANA es el organismo que controla la asignación de números para sistemas autónomos públicos que van en el rango de: 1 – 64511; los sistemas autónomos que se establecen como privados están en el siguiente rango: 64512 a 65535, estos ASN tienen una longitud de 2 bytes (16 bits) y proporcionan un máximo de 65535 ASN. Debido al rápido crecimiento de la tecnología estos

ASN se han ido agotando, por ello en el RFC 4893 se ha expandido la longitud de los ASN a 4 bytes (32 bits) y permiten un total de 4,294,967,295 ASN, y a partir del 4,200,000,000 al 4,294,967,294 se establecen como ASN privados el resto es de carácter público.

- Dentro de una misma red los sistemas autónomos hacen uso de protocolos de puerta de enlace interior (IGP). Y para la comunicación entre diversos sistemas autónomos utilizan EGP. La versión 4 de BGP es el único EGP actualmente en uso.
- Se le llama enrutamiento entre dominios al enrutamiento que se aplica entre sistemas autónomos.
- En cuanto a distancias administrativas eBGP tiene un valor de 20 mientras que iBGP de 200.
- Los vecinos BGP se denominan pares y deben configurarse estáticamente.
- Para el intercambio de actualizaciones de rutas incrementales, activadas y keepalives entre pares, BGP utiliza el puerto TCP 179.
- En los enrutadores solo se puede ejecutar una instancia de BGP a la vez.
- BGP utiliza el número de sistema autónomo como mecanismo para prevención de bucles. Es decir, cuando una actualización dentro una red sale de un sistema autónomo, el número de ese sistema autónomo se antepone a la lista de sistemas autónomos que han manejado esa actualización; en caso de encontrar su propio número de sistema autónomo en esa lista, la actualización se descarta.
- Dentro de una red iBGP se hace uso de una regla característica de BGP denominada Split-horizon, esta se usa para evitar bucles y consiste en que los prefijos aprendidos de un par iBGP nunca se anuncian a otro par iBGP (Bookham, 2014, p.263).
- A un enrutador que refleja las rutas se le conoce como Router Reflector (RR), esto ocurre en entornos iBGP; para prevenir los bucles en la operación de reflejar rutas los enrutadores deben operar bajo ciertas restricciones. Dentro de este enfoque los enrutadores iBGP se clasifican en tres grupos:
    - Route reflectors (RRs)
    - Route reflector clients (pares entre altavoces BGP)
    - Regular iBGP speakers (aquellos que no son clientes o no son pares BGP) (Zhang y Bartell, 2016, pp.254-255).

1.3.1.1. BGP Message Header Format

La Figura 4-1, muestra el formato de encabezado de un mensaje BGP.

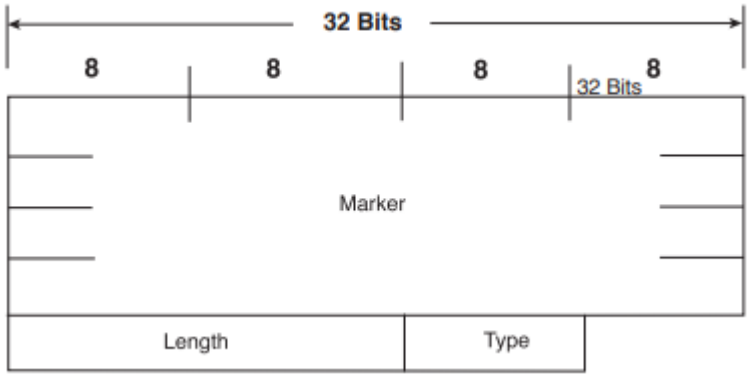

Figura 4-1: Encabezado de un mensaje BGP

Fuente: (Doyle, 2016, p.103)

- **Marker:** Este campo tiene una longitud de 16 bytes y su función es detectar si se ha perdido la sincronización entre pares BGP, además el campo marker sirve para el cálculo de la autenticación en BGP.
- **Length:** Es un campo de 2 bytes y expresa la longitud total del mensaje, incluyendo el encabezado.
- **Type:** Es un campo de 1 byte, este indica el tipo de mensaje que se envía, este campo posee 4 tipos de mensajes importantes que se describen en la tabla 2-1.

Tabla 2-1: Tipos de mensajes dentro del campo Type

CODIGO	TIPO	DESCRIPCION
1	OPEN	Después de establecer la sesión TCP, el mensaje open es el primero en ser enviado; cuando la recepción del mensaje open ha sido satisfactoria se envía como respuesta de confirmación un mensaje keepalive, tras esta exitosa acción los enrutadores proceden a enviar mensajes de actualización, mantenimiento y notificación. Dentro del mensaje open se encuentran campos importantes a saber: - Versión: Establece la versión de BGP, para BGP-4 el valor asignado es 4. - My Autonomous System: Se establece el número de AS del remitente. - Hold Time: Este campo indica el tiempo de espera en segundos establecido por el remitente del mensaje. Los pares BGP hacen uso de este campo para negociar el intervalo en el que se envían los mensajes Keepalive o Update con el fin de mantener la conexión entre ellos. El valor predeterminado es 180. - BGP Identifier: Establece el ID del enrutador remitente.

		- Optional Parameters: Contiene una lista de parámetros opcionales.
2	UPDATE	En este campo se anuncia cualquier ruta factible, retira rutas previamente anunciadas o puede hacer ambas cosas. El mensaje de actualización incluye la información de accesibilidad de la capa de red (NLRI) que incluye el prefijo y los PA de BGP asociados cuando se anuncian prefijos.
3	NOTIFICATION	Este campo se emplea para la notificación de errores entre pares de altavoces BGP, por ejemplo, cuando un temporizador expira, cuando se solicita el restablecimiento de una ruta o cuando las capacidades del vecino han cambiado.
4	KEEPALIVE	Los Keepalive se emplean para mantener las sesiones de peering entre altavoces BGP. Los mensajes Keepalive se intercambian por defecto cada 60 segundos.

Fuente: (Doyle, 2016, pp.103-108**;** Jain y Edgeworth, 2016, pp.6-8; HUAWEI, 2020)

Realizado por: Mena Duval, 2021

1.3.1.2. Sesiones BGP

Las adyacencias que se producen entre enrutadores configurados con BGP se conocen como sesiones BGP, estas sesiones son siempre punto a punto y se clasifican en dos tipos:

- **Internal BGP (IBGP):** Estas sesiones se establecen entre routers iBGP, es decir, que están en un mismo sistema autónomo (ASN) o que pertenecen a una misma confederación. IBGP se considera más seguro o en su defecto las medidas de seguridad pueden reducirse en comparación con eBGP.
- **External BGP (EBGP):** Son aquellas sesiones que se establecen entre enrutadores que pertenecen a diferentes sistemas autónomos (AS). (Jain y Edgeworth, 2016, pp.4-5)

1.3.1.3. BGP Path Selection

Para realizar una adecuada selección de ruta BGP asigna diversos atributos a cada ruta, estos atributos de ruta (PA) a su vez le permiten a BGP granularidad y control sobre las políticas de enrutamiento en el interior de una red configurada con BGP. Son cuatro las categorías existentes que organizan a los atributos de ruta, Donohue y Stewart (2010, p.76) explican dichas categorías de la siguiente forma:

- **Well-known mandatory (Conocido obligatorio):** Los campos de esta categoría deben ser reconocidos por todos los altavoces BGP, además tienen que estar presentes en todas las actualizaciones y transmitirse a todos los enrutadores configurados con BGP.

- **Well-known discretionary (Discrecional conocido):** Estos campos también deben ser reconocidos por los altavoces BGP y deben enviarse a los demás enrutadores BGP, pero no es obligatorio que estén en todas las actualizaciones.
- **Optional transitive (Opcional transitivo):** Los campos de esta categoría es posible que un altavoz BGP los reconozca o no, pero se transmite a otros enrutadores BGP. Si no se reconoce, se marca como parcial.
- **Optional nontransitive (Opcional no transitivo):** Los campos de esta categoría es posible que un enrutador BGP lo reconozca o no y que no se transmita a otros enrutadores.

Dentro de cada una de las categorías mencionadas anteriormente se encuentran los atributos de ruta (PA – Path Attributes), estos atributos influyen directamente a que BGP seleccione la mejor ruta y son fundamentales para el diseño de una arquitectura de red BGP eficiente. La tabla 3-1, muestra detalladamente cada uno de los atributos de ruta que forman parte del protocolo BGP.

Tabla 3-1: Atributos para la selección de ruta.

ATRIBUTO	CATEGORIA	DESCRIPCIÓN
ORIGIN	Well-known mandatory	En este atributo se indica el origen de la información de la ruta, son tres los posibles orígenes: - IGP: 0 - EGP: 1 - INCOMPLETE: 3 Para la selección de ruta se prefiere al prefijo que tener el valor más bajo.
AS_PATH	Well-known mandatory	El objetivo de este atributo es la prevención de bucles en el enrutamiento entre AS; el número de AS aceptados en esta lista es de 1 – 255; el número de AS se antepone como atributo para la selección de ruta.
NEXT_HOP	Well-known mandatory	Este atributo define la dirección IP del siguiente salto para alcanzar un prefijo desde el punto de vista de BGP. Esto no significa necesariamente que el siguiente salto esté conectado directamente; en este atributo se debe tomar en cuenta el TTL puesto que BGP lo establece en 1, es necesario modificar el TTL según el número de saltos que existen entre los ASN para lograr la comunicación.
LOCAL_PREF	Well-known discretionary	Un altavoz BGP usa este atributo para informar a sus otros pares internos del grado de preferencia del anunciante para una ruta anunciada.
ATOMIC_AGGREGATE	Well-known discretionary	Cuando un altavoz BGP agrega varias rutas con el propósito de anunciarlas a un peer en particular, el AS_PATH de la ruta agregada normalmente incluye un AS_SET formado por el conjunto de ases de los que se formó el agregado.

AGGREGATOR	Optional transitive	Este atributo contiene el último número AS que formó la ruta agregada (codificado en 2 octetos), seguido de la dirección IP del altavoz BGP que formó la ruta agregada (codificada en 4 octetos). Esta debería ser la misma dirección que la utilizada para el identificador BGP del altavoz.
COMMUNITY	Optional transitive	Este atributo se forma por un grupo de prefijos que comparten una propiedad en común. Dentro se un prefijo se pueden incorporar varias comunidades, estas comunidades tienen una longitud de 4 bytes y se dividen en dos: - Comunidades conocidas (Well-known communities) - Comunidades privadas (Private communities)
MULTI_EXIT_DISC (MED)	Optional non-transitive	Este atributo puede ser utilizado por el proceso de decisión de un hablante de BGP para discriminar entre múltiples puntos de entrada a un sistema autónomo vecino
ORIGINATOR_ID	Optional non-transitive	Este atributo se usa para prevenir los bucles entre enrutadores que están configurados como reflectores de ruta bajo un sistema autónomo (AS).
CLUSTER_LIST	Optional non-transitive	Este atributo también previene los bucles dentro de un AS que tiene configurados reflectores de ruta. Este atributo registra la lista de CLUSTER_ID que un prefijo ha atravesado en un entorno RR.

Fuente: (Zhang y Bartell, 2016**;** Rekhter et al., 2006)

Realizado por: Mena Duval, 2021

Puesto que BGP hace lo posible por seleccionar su mejor ruta, y lograr el objetivo de evitar bucles, siendo así que BGP se rige bajo ciertos criterios o parámetros de selección de ruta permitiéndole reducir el consumo de recursos y establecer la mejor ruta, a continuación Donohue y Stewart (2010, p.77) enumeran los criterios que dan soporte a BGP en su elección de ruta:

1. Ruta con mayor peso.
2. Ruta con la preferencia local más alta.
3. Elija las rutas que originó este enrutador.
4. Elija la ruta con la ruta más corta del sistema autónomo.
5. Elija la ruta con el código de origen más bajo
6. Elija la ruta con la MED más baja.
7. Elija una ruta EBGP sobre una ruta IBGP.
8. Elija la ruta a través del vecino IGP más cercano según lo determinado por la métrica IGP más baja.

9. Elige la ruta más antigua
10. Elija una ruta a través del vecino con la ID de enrutador más baja.
11. Elija una ruta a través del vecino con la dirección IP más baja.

1.3.1.4. Bases de Información de enrutamiento de BGP (RIBs)

BGP hace uso de tres tablas que almacenan información sobre encaminamiento, prefijos de red y los atributos de ruta (PA); Jain y Edgeworth (2016, p.17) brindan una explicación detallada del funcionamiento de las tres tablas:

- **Adj-RIB-in:** En esta tabla se almacena la información de accesibilidad de la capa de red (NLRI) antes de que se haya procesado las políticas de ruta de entrada, luego de este proceso la tabla se limpia con el objetivo de ahorrar memoria.
- **Loc-RIB:** En esta tabla están contenidos todos los NLRI que fueron originados localmente o que se han recibido de otros pares de altavoces BGP. Una vez que estos NLRI pasen la comprobación de validez y alcanzabilidad, el algoritmo de mejor ruta selecciona el mejor NLRI para un prefijo especifico; esta tabla presenta las rutas a la tabla de enrutamiento IP.
- **Adj-RIB-out:** Contiene los NLRI después de que se hayan procesado las políticas de ruta de salida. Esta tabla se confecciona a raíz de las informaciones de la tabla Loc-RIB.

1.3.1.5. BGP Neighbor States

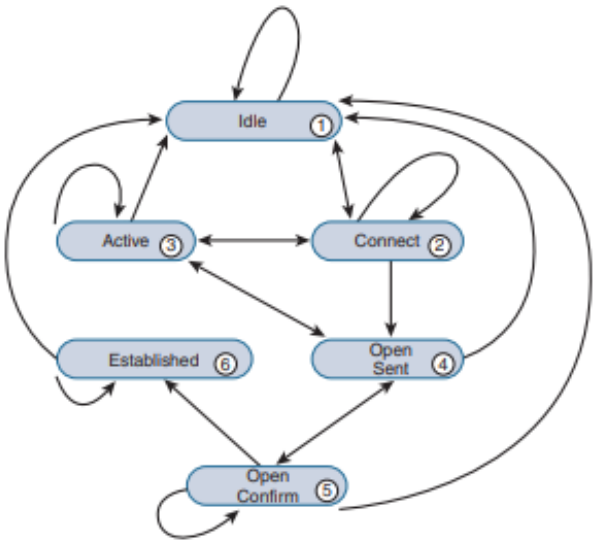

Figura 5-1: Máquina de estados finito BGP

Fuente: (Jain y Edgeworth, 2016, p.8)

Al establecer sesiones TCP con los peers vecinos, BGP procede a confeccionar una tabla con los estados posibles de un vecino BGP, esto lo logra haciendo uso de una herramienta llamada máquina de estados finito (FSM), en la tabla se almacenará información de los peers BGP y su estado operativo. La figura 5-1 muestra la forma en cómo funciona la máquina de estados finito (FSM) y todos sus estados posibles. Cada estado representa un acontecimiento especifico asociado al vecino BGP, Jain y Edgeworth (2016, pp.8-10) explican cada estado a continuación:

- **Idle:** Es la primera operación de la máquina de estados BGP, en este estado BGP intenta realizar una conexión TCP con el peer BGP luego de detectar un evento de inicio.
- **Connect:** En este estado se establece la sesión TCP, una vez que el proceso 3-way handshake se ha completado, BGP reinicia el ConnectRetryTimer y procede a enviar el mensaje open al vecino, acto seguido cambia al estado OpenSent.
- **Active:** En este estado, BGP inicia un nuevo proceso 3-way handshake. Al establecer conexión se envía el mensaje open al vecino, se establece el Hold timer en 4 minutos, posteriormente se pasa al estado OpenSent; si esta conexión falla el estado vuelve a connect y se reinicia el ConnectRetryTimer.
- **OpenSent:** Una vez llegado a este estado, el enrutador de origen envía un mensaje open a su vez esperando un mensaje open del router receptor, aquí se comparan los errores de ambos mensajes open teniendo en cuenta los siguientes elementos:
    - La versión de BGP debe ser la misma en los pares BGP.
    - La dirección IP del open de origen debe coincidir con la dirección IP que está configurada para el vecino.
    - El número de AS del mensaje OPEN debe coincidir con el configurado para el vecino.
    - Los identificadores BGP (RID) deben ser únicos.
- **OpenConfirm:** Este estado se refiere al recibo de un mensaje de confirmación, es decir, al recibir un keepalive o un mensaje de notificación la FSM pasa al estado de establecido; si el temporizador expira se vuelve al estado de idle (inactivo).
- **Established:** En este estado los altavoces BGP envían mensajes de actualización, keepalives e intercambian rutas entre pares vecinos puesto que ya se establecido de manera correcta la sesión TCP. Si el temporizador hold time expira la maquina FSM vuelve al vecino al estado de inactivo (idle).

1.3.2. Multiprotocol BGP (MP-BGP)

Dentro del RFC 4760 se define el multiprotocolo BGP (MP-BGP), dicho RFC describe que el objetivo principal del MP-BGP es permitir una multiplexación de familia de direcciones y NLRI acoplado en una sola sesión de peering BGP; además BGP incluye servicios VPN a través de esta única sesión (Krattiger et al., 2017). El Multiprotocolo BGP es normalmente abreviado como MBGP, M-BGP o MP-BGP, en adelante se usará el acrónimo MP-BGP ya que se considera que es el más representativo.

Para permitir la simplificación en la introducción de las capacidades multiprotocolo y proporcionar compatibilidad con versiones anteriores, BGP introduce dos nuevos conceptos con respecto a los atributos de BGP: Multiprotocol Reachable NLRI (MP_REACH_NLRI) y Multiprotocol Unreachable NLRI (MP_UNREACH_NLRI). El atributo (MP_REACH_NLRI) es utilizado para transportar el conjunto de destinos alcanzables más la información del siguiente salto que se utilizará para el reenvío a estos destinos; en cambio el atributo (MP_UNREACH_NLRI) se utiliza para llevar el conjunto de destinos inalcanzables. Estos atributos son opcionales y no transitivos (Optional non-transitive), lo que significa que si un altavoz BGP no tiene soporte para las funcionalidades del multiprotocolo BGP simplemente va a ignorar la información recibida por estos atributos y tampoco compartirá la información con los demás altavoces BGP (Bates et al., 2007, p.2).

Tabla 4-1: Familias de direcciones comúnmente usadas en BGP con sus respectivos identificadores para AFI/SAFI.

AFI	SAFI	Network Layer Information
1	1	IPv4 Unicast
1	2	IPv4 Multicast
1	4	IPv4 Unicast with MPLS Label
1	128	MPLS L3VPN IPv4
2	1	IPv6 Unicast
2	4	IPv6 Unicast with MPLS Label
2	128	MPLS L3VPN IPv6
25	65	Virtual Private LAN Service (VPLS) Virtual Private Wire Service (VPWS)
25	70	Ethernet VPN (EVPN)

Fuente: (Jain y Edgeworth, 2016, p.4)

Realizado por: Mena Duval, 2021

Dentro del MP-BGP los protocolos que son admitidos por BGP poseen un mecanismo que ayuda a identificarlos a todos y cada uno, este es conocido como identificador de familia de direcciones

(AFI). AFI se encarga de identificar el protocolo principal, por ejemplo, IPv4 o IPv6; sin embargo, esto no es suficiente para poder identificar de forma asertiva a todos los procesos que incluye BGP, por lo tanto, es necesario más distinciones. Por ejemplo, la información de accesibilidad unidifusión y multidifusión difiere significativamente, es así que BGP distingue mediante el uso de números de Indicador de familia de direcciones subsiguientes (SAFI) para reconocer entre las direcciones de unidifusión y multidifusión. IPv4 Unicast AFI-SAFI es el que se toma como base en el RFC. Por lo tanto, cuando no se especifica AFI / SAFI, se supone que se aplica a IPv4 Unicast. La lista AFI / SAFI que es de interés para un altavoz BGP se anuncia utilizando las capacidades BGP en el mensaje OPEN; se logra el intercambio de información solo si los pares BGP anuncian su interés por dicho AFI/SAFI (Dutt, 2019, p.293).

Cada familia de direcciones proporciona una base de datos y una configuración independiente para cada protocolo (AFI + SAFI) dentro de BGP, esto permite lograr una diferenciación entre las familias de direcciones en MP-BGP. BGP incluye un AFI y un SAFI con cada anuncio de ruta para diferenciar entre las bases de datos AFI y SAFI. La tabla 4-1, muestra algunas de las familias de direcciones AFI/SAFI utilizadas comúnmente en MP-BGP (Jain y Edgeworth, 2016, p.4).

1.3.2.1. Multiprotocolo alcanzable NLRI (MP_REACH_NLRI)

El atributo MP_REACH_NLRI es transportado dentro de los mensajes de actualización (Update) con el objetivo de anunciar rutas accesibles; la función de MP_REACH_NLRI es identificar el protocolo que se anuncia y el tipo de dirección correspondiente al protocolo, el siguiente salto de la ruta y el propio NLRI. La Tabla 5-1 muestra los campos contenidos dentro del MP_REACH_NLRI (Doyle, 2016, p.616).

Tabla 5-1: Campos del atributo MP_REACH_NLRI.

Address Family Identifier (2 octets)
Subsequent Address Family Identifier (1 octet)
Length of Next Hop Network Address (1 octet)
Network Address of Next Hop (variable)
Reserved (1 octet)
Network Layer Reachability Information (variable)

Fuente: (Bates et al., 2007, p.3)

Realizado por: Mena Duval, 2021

A continuación, Doyle (2016, pp.616-617) explica el uso y significado de cada uno de los campos contenidos en el atributo MP_REACH_NLRI.

- **Address Family Identifier (AFI):** La función de este campo es identificar y especificar el protocolo al que pertenece el prefijo NLRI, es una larga lista los números de protocolos correspondientes a AFI por ejemplo para IPv4 es AFI 1 y para IPv6 es AFI 2.
- **Subsequent Address Family Identifier (SAFI):** En este campo se especifica un tipo de dirección funcional según el protocolo principal, por ejemplo, para unicast NLRI es SAFI 1 y para multicast NLRI es SAFI 2.
- **Next Hop Address Length:** En este campo se especifica la longitud de la dirección de siguiente salto (Next-hop) del protocolo. La longitud máxima que se puede especificar es 256 bytes, se considera más que suficiente para cualquier dirección ya que para una dirección IPv4 solo se necesita 4 bytes y para una IPv6 se necesita 16 bytes; hay ocasiones que se ocupa dos direcciones IPv6 global y local en el campo de siguiente salto lo que significaría un total de 32 bytes que siguen siendo suficientes para poder trabajar dentro de los 256 bytes.
- **Next Hop Address:** Aquí está contenida la dirección de la capa de red, que cumple con el formato de dirección AFI/SAFI especificados para el siguiente salto, correspondiente al NLRI anunciado.
- **Network Layer Reachability Information (NLRI):** Aquí está contenido el prefijo y la longitud del prefijo correspondiente al formato de dirección AFI/SAFI especificado. Dentro de IP el prefijo/longitud es equivalente a la notación CIDR para un prefijo IPv4 o IPv6.

1.3.2.2. Multiprotocolo inalcanzable NLRI (MP_UNREACH_NLRI)

El atributo MP_UNREACH_NLRI se utiliza para descartar del servicio múltiples rutas no viables, este atributo es opcional y no transitivo (Optional non-transitive). La tabla 6-1 muestra los campos contenidos dentro del atributo MP_UNREACH_NLRI.

Tabla 6-1: Campos del atributo MP_UNREACH_NLRI

Address Family Identifier (2 octets)
Subsequent Address Family Identifier (1 octet)
Withdrawn Routes (variable)

Fuente: (Bates et al., 2007, p.5)

Realizado por: Mena Duval, 2021

El RFC 4760 realizado por el grupo de trabajo de la IETF, contiene la descripción de todos los campos que forman parte del atributo MP_UNREACH_NLRI:

- **AFI/SAFI:** La combinación entre el Identificador de la familia de direcciones (AFI) y el Identificador de la familia de direcciones subsiguiente (SAFI), permiten identificar el conjunto de protocolos de la capa de red al que pertenece la dirección que es transportada en el campo NEXT-HOP, además se encuentra la forma en cómo se codifica la dirección del campo Next-hop y la semántica de la información de alcanzabilidad de la Capa de Red (NLRI) (Bates et al., 2007, pp.5-6).
- **Withdrawn Routes NLRI:** La longitud de este campo es variable y tiene por objetivo enumerar los NLRI de las rutas que se retiran del servicio.

1.3.2.3. Capacidades BGP (Capabilities)

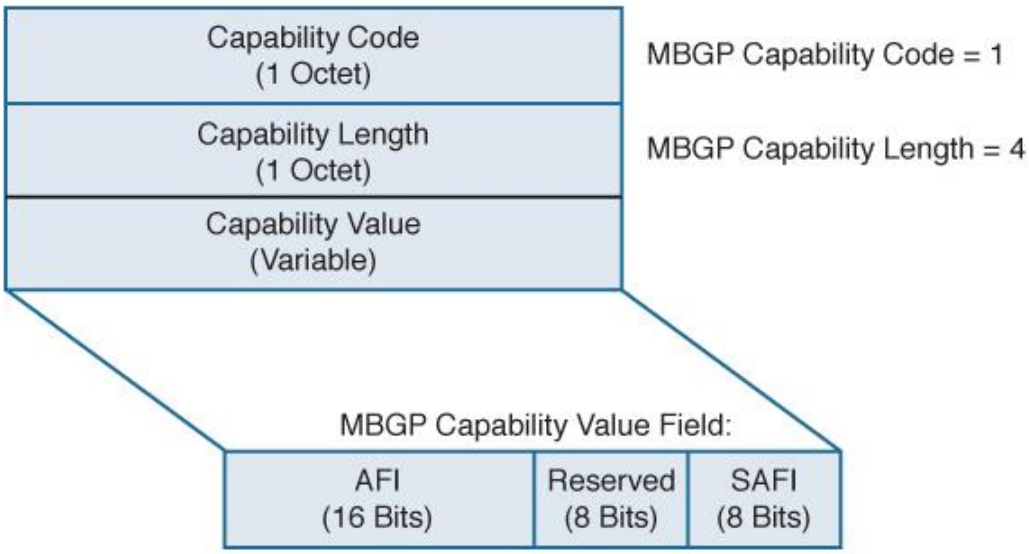

Figura 6-1: Opciones de capacidades MP-BGP

Fuente: (Doyle, 2016, p.618)

MP-BGP posee una función importante que le da soporte denominada capacidades BGP, esta se encarga de brindar capacidad para que un altavoz BGP pueda incluir en sus mensajes OPEN, una descripción de las capacidades opcionales que admite. Es decir, cuando un vecino ha establecido adyacencia y recibe un mensaje OPEN con el parámetro capacidades, este puede aceptar dichas capacidades o enviar un mensaje de error tipo 7 indicando que no admite una o más capacidades. Esta notificación se hace uso solo cuando el vecino tiene conocimiento sobre estas capacidades caso contrario si el vecino no tiene ningún conocimiento sobre estas capacidades no envía ninguna notificación (Doyle, 2016, p.617).

La figura 6-1, muestra el formato del parámetro capacidades, además de la configuración que el parámetro utiliza para indicar la compatibilidad con el MP-BGP. Para distinguir el uso de capacidades dentro del multiprotocolo BGP se asigna el valor 1 en Capability code y el valor 4 en el campo Capability Length hace referencia a la longitud en octetos del valor de capacidad (Doyle, 2016, p.618).

1.4. Virtual Extensible Local Area Network (VXLAN)

1.4.1. Introducción

Debido al crecimiento exponencial de la tecnología de virtualización, los administradores de centros de datos han tenido que acoplarse a las exigencias de los usuarios y aplicaciones, albergando un mayor número de servidores por bastidor e incrementando las estructuras físicas de sus centros de datos; pero siempre quedará algo que mejorar para poder seguir creciendo, es el caso de STP y VLANs, puesto que STP previene bucles a costa del bloqueo de enlaces y reducción del ancho de banda, sumado a esto el uso de VLANs que en su momento tenía una gran acogida pero pronto la tecnología establecería que 4096 vlans no son suficiente para escalar un centro de datos.

Es por dichas limitaciones que los ingenieros de redes han tenido la necesidad de encontrar soluciones alternativas; pero se debe tomar en cuenta que preservar el uso de la semántica del puente de OSI (L2) para establecer comunicación entre máquinas virtuales y dispositivos heredados dentro de las redes es necesario. Es por esto que se desarrolló VXLAN (Virtual Extensible Local Area Network). VXLAN permite la ampliación de 4096 ID de VLAN (12 bits) a 16 millones por medio de su VNI (Identificador de red VXLAN – 24 bits). VXLAN representa una superposición de capa 2 sobre capa 3, permitiendo conectar dos o más redes a través de la capa 3 pero manteniendo las cargas de trabajo de los servidores bajo el dominio de transmisión de capa 2, por ello, mientras que VLANs sigue trabajando dentro de la capa 2, VXLAN opera en la capa 3. Y recordando a STP con sus múltiples desventajas junto con las VLANs, tenemos que, VXLAN hace uso del protocolo ECMP que permite aprovechar todas las rutas disponibles para el transporte de información (Krattiger et al., 2017).

Para lograr el descubrimiento de los VNI en la red subyacente de Capa 3, se utilizan puntos finales de túnel virtual (VTEP). A los VTEP se les considera como una entidad que termina los túneles VxLAN. Estos operan asignando tramas de capa 2 a un VNI para su uso en la red de superposición. La encapsulación del tráfico de Capa 2 y Capa 3 del cliente en VNI a través de la red física de Capa 3 proporciona el desacoplamiento de la superposición de la red subyacente y

proporciona una topología de superposición flexible que es independiente de la topología de la red física (Jain y Edgeworth, 2016, pp.641-642). Los VTEP cuentan con dos interfaces.

- **Local LAN segment:** Dentro de la red local el VTEP permite una funcionalidad de puenteo, es decir, los hosts conectados en el segmento LAN pueden lograr la comunicación a través de puentes administrados por el VTEP.
- **IP interface:** Esta dirección IP de la interfaz es la que permite identificar de forma única a los VTEPs en la red, además aquí se realiza el proceso de encapsulación/des encapsulación de VXLAN (Jain y Edgeworth, 2016, p.642).

La figura 7-1, muestra el esquema que utilizan los VTEP, además de como los hosts finales se conectan a los mismos.

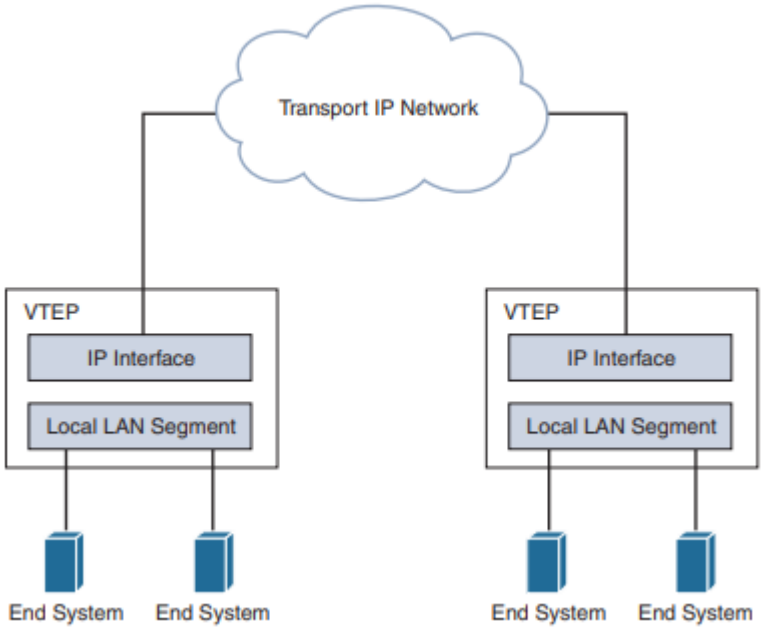

Figura 7-1: Esquema del VTEP y forma de conexión de los sistemas finales.

Fuente: (Jain y Edgeworth, 2016, p.642)

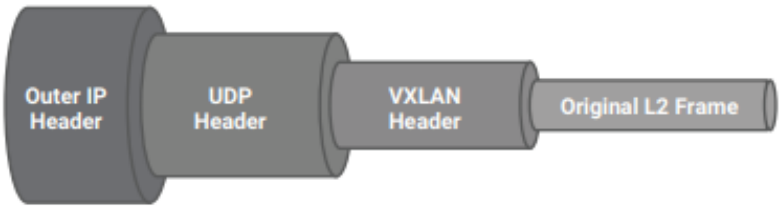

Figura 8-1: Formato de encapsulación VXLAN.

Fuente: (Gai, 2020, p.20)

En la figura 8-1 se puede apreciar el formato de encapsulación de VXLAN. Como se dijo VXLAN realiza una encapsulación de MAC sobre IP/UDP; el puerto de destino de UDP se establece en el valor 4789 mientras que el puerto de origen es generado aleatoriamente en función del encabezado original o externo.

A continuación, se describen dos ejemplos de cómo se generaría el puerto de origen UDP para VXLAN: En el primer ejemplo el VTEP puede configurar el puerto UDP de origen para que sea igual a un hash de la dirección IP de origen de la trama ethernet original, es decir, del encabezado IP interno que normalmente pertenece a una VM; esta primera opción implementa el equilibrio de carga a nivel de VM. El segundo ejemplo se basa en que el VTEP puede generar el puerto de origen en un hash a partir del mecanismo de tupla-5 que normalmente se compone de la dirección IP de origen, la dirección IP de destino, el protocolo de Capa 4, el puerto de origen de L4 y el puerto de destino de L4 (Gai, 2020, p.20).

1.4.2. Formato de la trama VXLAN

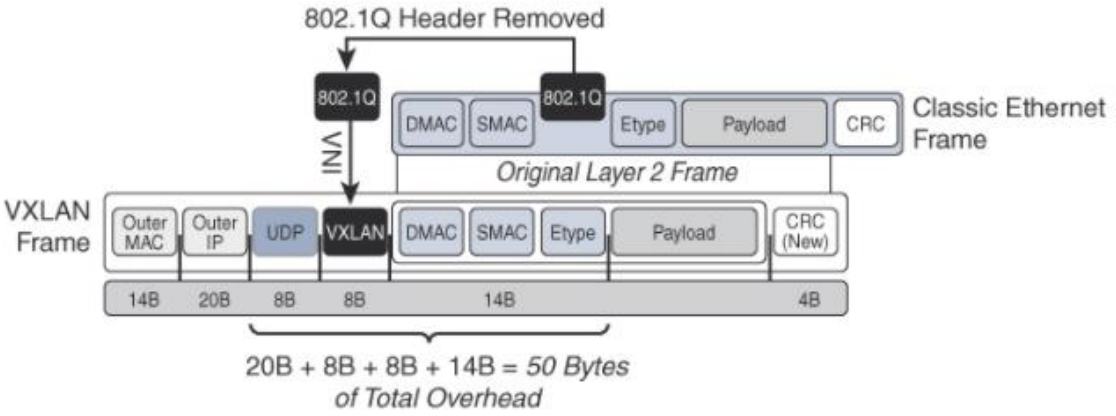

Figura 9-1: Formato de trama VXLAN.

Fuente: (Krattiger et al., 2017)

En el encabezado VXLAN los componentes internos se refieren a la capa 2, mientras que los componentes externos hacen referencia a la capa 3. Es así que VXLAN contiene en su interior la trama Ethernet de capa 2 original como carga de trabajo, que a su vez está compuesta por la dirección MAC de origen interna, dirección MAC de destino interna, la carga útil Ethernet original y la secuencia de verificación de tramas (FCS). En la figura 9-1, se puede apreciar que en el encabezado interno de VXLAN la etiqueta 802.1Q de la trama ethernet original de capa 2 es eliminada y asignada a un VNI con el objetivo de completar la trama del encabezado VXLAN (Krattiger et al., 2017).

A continuación la figura 10-1, muestra detalladamente los campos contenidos dentro de la trama VXLAN, Jain y Edgeworth (2016, p.643) describen los campos representativos de la cabecera VXLAN:

- **Flags:** Las banderas son un campo de 8 bits de longitud, por lo tanto, son 8 banderas, donde 7 de estas se establecen como reservadas (bits R) y obtienen el valor 0; sin embargo, la quinta bandera (bandera I) se le asigna el valor de 1, lo que hace referencia a que existe un VNI válido.
- **VNI:** Este valor es el que permite la identificación única para cada segmento de VXLAN, este campo posee una longitud de 24 bits.

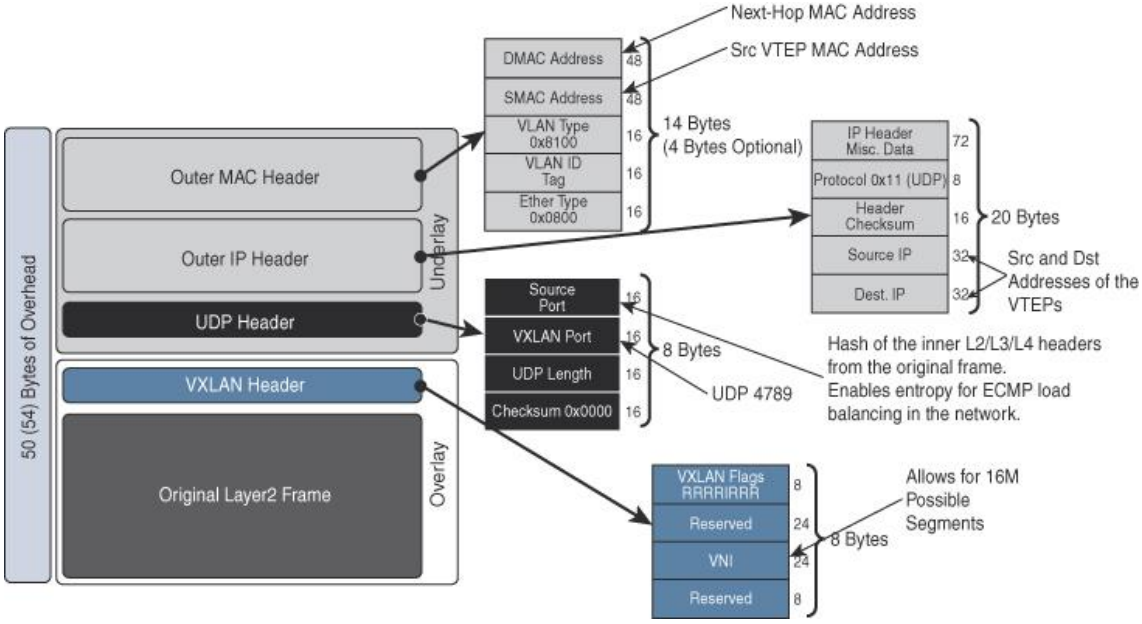

Figura 10-1: Detalles del formato de la trama VXLAN

Fuente: (Krattiger et al., 2017)

En el RFC 7348 elaborado por la IETF en colaboración con Mahalingam et al. (2014, pp.10-11), se detallan los campos que hacen referencia a la base (underlay), también considerados como campos exteriores, que se muestran en la figura 10-1, pertenecientes al formato de la trama VXLAN:

- **Outer UDP Header:** El puerto de origen es establecido por el VTEP y se calcula utilizando el hash de tupla-5 o un hash relacionado con el encabezado IP interno, esto se realiza con el objetivo de tener una entropía con ECMP para el equilibrio de carga del tráfico VM a VM. El número del puerto de destino está establecido en 4789 y fue asignado por la IANA, además este puerto debe ser configurable.
- **Outer IP Header:** La dirección IP de origen en el encabezado IP externo de la trama VXLAN, es la interfaz IP del VTEP de origen. Esta dirección IP es la que identifica de forma

exclusiva a un VTEP. La dirección de destino del encabezado IP externo es la dirección IP de la interfaz IP del VTEP de destino.

- **Outer Ethernet/MAC Header:** la dirección MAC de origen es la dirección MAC del VTEP de origen. La dirección MAC de destino es la dirección MAC del siguiente salto. El siguiente salto es la interfaz utilizada para llegar al destino o al VTEP remoto (Jain y Edgeworth, 2016, p.644).

1.4.3. Tipos de puerta de enlace VXLAN

El VTEP es aquel que se encarga de realizar la encapsulación y desencapsulación de la trama en VXLAN, de esta manera los VTEP inician y terminan el túnel VXLAN; la conexión del tráfico entre un segmento VXLAN y otro dominio de capa 2 se logra a través de una puerta de enlace (Gateway VXLAN). Existen dos tipos de puertas de enlace para VXLAN, Jain y Edgeworth (2016,p.645) los explica a continuación:

- **Layer 2 Gateway:** Se hace uso de la puerta de enlace de capa 2, cuando el tráfico de capa 2 que está etiquetado con IEEE 802.1Q proviene de la VLAN y se dirige hacia un segmento VXLAN encapsulado; o a su vez, el paquete de entrada VXLAN sale de una interfaz etiquetada con IEEE 802.1Q desencapsulado, para posteriormente puentear el paquete a una nueva VLAN.
- **Layer 3 Gateway:** Gateway L3 se utiliza cuando existe un enrutamiento de VXLAN a VLAN, es decir, cuando el paquete VXLAN de salida se enruta a un nuevo segmento VXLAN. Además, se utiliza la puerta de enlace L3 cuando existe enrutamiento entre VXLAN hacia una VLAN, es decir, cuando el paquete de entrada es VXLAN en un segmento enrutado y va de salida hacia una interfaz 802.1q etiquetada, para posteriormente volver a enrutar el paquete hacia una nueva VLAN.

1.4.4. Variantes para la implementación de VXLAN

El mecanismo por el cual las entidades forman relaciones para el intercambio de información dentro de una red, es definido bajo un protocolo de plano de control, puesto que influye directamente en el plano de datos ya que se interrelacionan entre sí para formar un mecanismo sólido y eficiente. VXLAN como tal, no posee un plano de control definido, por lo tanto, el mecanismo que se encarga del intercambio de información o el aprendizaje de direcciones es el plano de datos y lo logra basándose en el mecanismo de inundación y aprendizaje (flood & learn), típico de capa 2. De forma que cuando un VTEP aprende la dirección MAC de un host final local, crea una entrada en su tabla de reenvió; y al no tener un plano de control definido, las direcciones

aprendidas no se anuncian a los demás VTEPs remotos, esto crea la necesidad de que cada VTEP tenga que realizar el aprendizaje de direcciones de forma individual en toda la red. Cuando un VTEP recibe un paquete VXLAN (aprendizaje remoto), crea una asociación donde se almacena la MAC remota, su respectivo VNI y el VTEP remoto, esto se hace con el objetivo de poder tener un mapeo de estos valores y saber cómo reenviar el tráfico destinado a esta MAC remota en el futuro. Para el tráfico multidestino (Broadcast, Unknown Unicast y Multicast), VXLAN recurre a un mecanismo de flooding (Chandra, 2017, pp.15-16).

Para la superposición VXLAN es necesario que todos los VTEPs se emparejen entre si para que la información pueda llegar correctamente a todos los destinos, por ello existe algunas variantes que VXLAN utiliza para habilitar el aprendizaje de MAC y manejar el tráfico BUM, Chandra (2017, pp.15-16) detalla dichas variantes de la siguiente forma:

- **VXLAN/Multicast - Flood and Learn (F&L):** El plano de datos es el que permite el aprendizaje MAC remoto. En esta variante se hace uso de grupos de multidifusión (multicast) ya que al agrupar diferentes segmentos VXLAN en un mismo grupo de multidifusión, se logra una limitación en el tráfico BUM, es decir el tráfico de capa 2 se envía solo a los VTEPs interesados. Cada segmento VXLAN es asignado a un grupo de multidifusión determinado. Para el caso de tráfico unicast unknown o broadcast, el VTEP inundará el paquete VXLAN enviándolo sólo al grupo multicast asociado al VNI o segmento VXLAN dado. Esta variante de multidifusión es conocido como grupos de multidifusión PIM y para que VXLAN pueda hacer uso de la misma, debe estar configurada debidamente en la capa base (subyacente). Con la replicación funcionando en la base, los paquetes se replican en el punto más lejano de la topología física, por lo que es eficiente para el tráfico multicast pesado. El inconveniente en esta variante es que muchos administradores de red no adoptan la configuración de grupos de multidifusión ya que supone una importante sobrecarga operativa, sumado a esto en términos de eficiencia en la red es todo lo contrario a facilitar la escalabilidad, puesto que si bien es cierto que VXLAN puede soportar 16 millones de redes lógicas de capa 2, no es posible soportar un mapeo 1:1 entre tantos VNIs y grupos de multidifusión, ya que esto se traduce a necesitar una cantidad descomunal de recursos de software y hardware, por lo que su implementación no es práctica.

- **VXLAN/Unicast - Ingress Replication:** En esta variante también el aprendizaje MAC remoto lo realiza el plano de datos. Mediante la utilización de la replicación de ingreso (Ingress Replication), el VTEP de origen replicará un paquete BUM localmente y enviará una copia unicast a cada uno de los otros VTEPs que participan en el mismo segmento VXLAN. Se pueden configurar múltiples VTEPs unicast para un segmento VNI determinado. Si bien es cierto esta configuración es más sencilla de implementar y gestionar, además de

proporcionar un control granular sobre el alcance de las inundaciones (flooding), inhabilita el autodescubrimiento de los cambios dinámicos de la topología. Sumado a esto, el procesamiento adicional en el nodo principal podría no hacer de esta una opción ideal para el tráfico multicast pesado.

- **OVSDB/VXLAN:** Esta variante es un enfoque a SDN (redes definidas por software), aquí la forma en que se realiza el aprendizaje MAC remoto es por medio del plano de control; normalmente OVSDB se utiliza junto con un controlador de superposición. Donde la replicación del tráfico BUM es administrada por el controlador de superposición adoptado, OVSDB hace uso de este controlador con la finalidad de habilitar el aprendizaje de direcciones MAC y gestiona el aprovisionamiento VTEP.
- **EVPN/VXLAN:** En esta variante EVPN (Ethernet VPN) es el encargado del plano de control que se combina junto con MP-BGP para lograr el objetivo del aprendizaje MAC remoto y anunciar la accesibilidad de direcciones, en el presente trabajo esta es la variante que se ha seleccionado para poner en marcha la superposición VXLAN, en el apartado 1.5., se explica con mayor detalle el funcionamiento de EVPN/VXLAN.

1.4.5. MTU Consideraciones

Se debe tomar en consideración que VXLAN como una tecnología de superposición agrega una sobrecarga a la trama ethernet original, esta sobrecarga contempla diferentes campos necesarios para el correcto funcionamiento de VXLAN, puesto que ayuda a superar varias fallas en el transporte de redes tradicionales (como Ethernet), por lo tanto, al configurar la base (underlay) sobre la cual funcionará VXLAN se debe tomar las consideraciones necesarias para evitar problemas a futuro. Como una regla general se debe evitar la fragmentación de la información, puesto que la fragmentación y el reensamblaje de los paquetes producen un mayor consumo de recursos en los conmutadores y afecta notablemente en el rendimiento de la red (Krattiger et al., 2017).

Históricamente, el protocolo IPv4 tenía la capacidad de fragmentar el paquete cuando se excedía el valor establecido de MTU, pero los enrutadores IPv4 modernos generalmente no fragmentan los paquetes IP; los descartan si exceden la MTU del enlace, lo que aumenta la dificultad para solucionar problemas de funcionamiento. Cabe mencionar que en IPv6 no se admite la fragmentación (Gai, 2020, p.22).

En la figura 10-1, se puede apreciar que la sobrecarga VXLAN lleva consigo los siguientes campos: encabezado MAC externo de 14 bytes, un encabezado IP externo de 20 bytes, un encabezado UDP de 8 bytes y un encabezado VXLAN de 8 bytes. La sobrecarga da como

resultado un total de 50 bytes, además existe una variante de 54 bytes, estos 4 bytes resultan de preservar la etiqueta IEEE 802.1Q en la trama Ethernet original. En la mayoría de casos la etiqueta IEEE 802.1Q es eliminada y reemplazada por el VNI de VXLAN, pero existen casos especiales como hipervisor anidado, tunelización de capa 2, que requieren conservar la etiqueta IEEE 802.1Q original, la conservación de la etiqueta que da una sobrecarga de 54 bytes en VXLAN se muestra en la figura 11-1 (Krattiger et al., 2017).

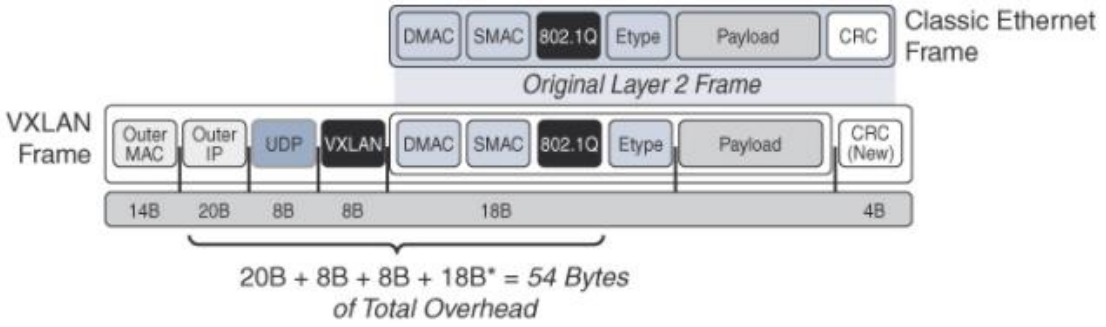

Figura 11-1: Sobrecarga de 54 Bytes en VXLAN

Fuente: (Krattiger et al., 2017)

De esta forma para cumplir con la regla de evitar la fragmentación, se debe hacer un pequeño ajuste en función del MTU, en este caso se debería configurar una MTU de 1550 o 1554 si se conserva la etiqueta IEEE 802.1Q original, las tramas que sobrepasan los 1500 bytes se consideran como tramas gigantes (jumbo frames), por ello los ingenieros de redes deben tomar en consideración el uso obligatorio de tramas gigantes en la red, por lo tanto, con estas consideraciones el problema de la sobrecarga se vería resuelto, sin embargo, una buena recomendación y practica por parte de VMware (2021, p.68) es configurar la interfaz con una MTU de 1600, es decir 100 bytes adicionales a la trama ethernet, puesto que así se tomaría en consideración y en forma de solución, la integración a futuro de protocolos adicionales que podrían aumentar la sobrecarga de trabajo y terminen causando problemas en la red.

Ahora bien, el enfoque del presente trabajo se centra en la aplicación de VXLAN/EVPN en centros de datos, por lo que la eficiencia es un aspecto importante, por ello una práctica común es ajustar las jumbo frames de lado del servidor ya que su uso mejora la eficiencia y, en una red de alta velocidad, se logra reducir el número de viajes de ida y vuelta necesarios a través de la red; estas tramas normalmente pueden tener un máximo de 9000 bytes, que es lo típico que ofrecen la gran mayoría de las NIC y/o conmutadores virtuales, de este modo si se permite una trama gigante de 9000 bytes por parte del servidor, se debe tomar en cuenta la sobrecarga de VXLAN y ajustar la MTU a 9050 o 9054 bytes, sin embargo para efectos de este trabajo se ha tomado la forma de configuración mencionada en VMware (2021, p.68), y se han ajustado las interfaces con una MTU

de 9100 bytes ya que se considera una buena práctica y varios ingenieros de redes la han adoptado (Krattiger et al., 2017).

1.5. VXLAN BGP EVPN

La sección anterior cubrió uno de los temas más importantes de este trabajo VXLAN. VXLAN ha sido considerada como la solución para la superposición y el derroque de STP que alejaba a los centro de datos de ser escalables, sin embargo, VXLAN por sí solo tiene dificultades como muchos otros protocolos, es muy complicado pretender embutir todas las necesidades de un centro de datos bajo una sola tecnología, es por esto que VXLAN hace uso de EVPN ya que le brinda el soporte en cuanto al plano de control, a su vez EVPN usa como cerebro de procesamiento a MP-BGP, para con este tridente lograr el objetivo de extender de forma eficiente la capa 2 sobre la capa 3.

1.5.1. Ethernet VPN (EVPN)

EVPN es una tecnología que permite interconectar segmentos de red L2 separados por una red L3. EVPN logra esto mediante la construcción de la red L2 como una superposición de red virtual de Capa 2 sobre la red de Capa 3. El cerebro de EVPN es el MP-BGP y lo utiliza como protocolo de control y en el centro de datos utiliza VXLAN para la encapsulación de los paquetes (Dutt, 2019, p.335). EVPN se considera como una alternativa un tanto compleja, que nació debido a la necesidad de superar las limitantes que estaban presentes en el servicio de LAN privada virtual (VPLS). EVPN cumple dos funciones: en primera instancia como ya se ha mencionado, facilita un plano de control de aprendizaje MAC para redes superpuestas, y la segunda es la necesidad de movilidad de las cargas de trabajo. Puesto que, las aplicaciones y cargas de trabajo necesitan del dominio de capa 2 para poder comunicarse entre sí; esto es algo normal dentro de un mismo centro de datos, la raíz del problema se genera cuando dichas aplicaciones y cargas de trabajo necesitan comunicarse entre dos o más centros de datos. Esto se traduce en extender las VLANs a través de la WAN, es ahí donde entra en juego el papel de EVPN (Wrightson, 2016, p.103).

Anteriormente VPLS era el encargado de extender el dominio de capa 2 a diferentes sitios, y de hecho realizaba un buen trabajo; sin embargo, como cualquier otro protocolo venia acarreando varias limitaciones en cuanto a: Escalado de direcciones MAC, soporte para multidifusión de forma sensata, multi-homing activo/activo, transporte transparente de direcciones MAC de cliente, convergencia más rápida, y sin duda el mayor dolor, la facilidad de gestión. Es así que EVPN intenta cubrir todos estos problemas (Wrightson, 2016, p.104).

En cuanto al plano de datos EVPN tiene múltiples tecnologías que son compatibles. La figura 12-1, muestra las tecnologías que EVPN puede utilizar para realizar las funciones del plano de datos, cabe mencionar que para efectos de este trabajo simplemente se tomará en cuenta el trabajo del plano de datos entre EVPN/VXLAN descrito en el borrador draft-ietf-bess-evpn-overlay.

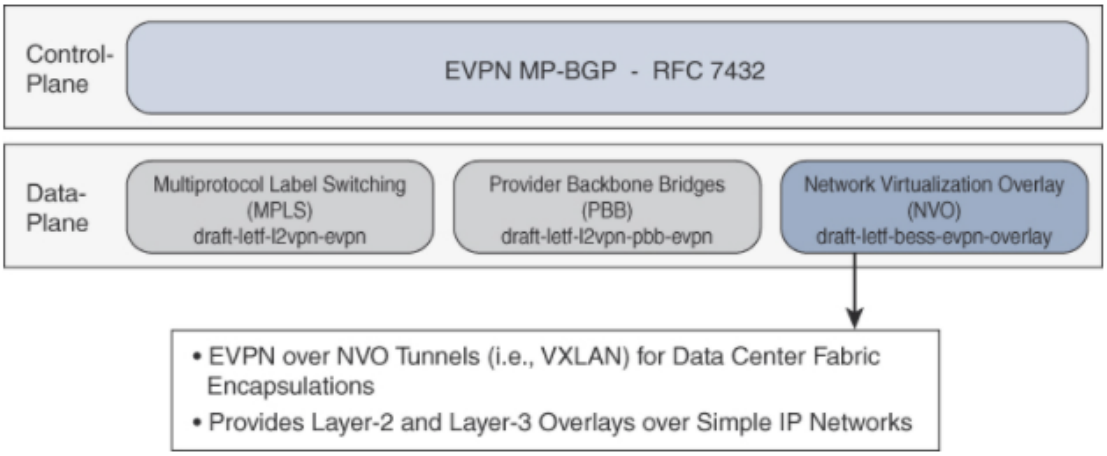

Figura 12-1: Soporte de plano de datos de EVPN

Fuente: (Krattiger et al., 2017)

EVPN introduce una nueva terminología específica para un entorno NVO, el grupo de trabajo de Arista Networks (2021), resume dicha terminología en relación con la encapsulación VXLAN:

- **Superposición de virtualización de red (NVO):** Hace referencia a la red superpuesta que se utiliza para ofrecer los servicios VPN de capa 2 y capa 3. En la encapsulación VXLAN, esto definiría un dominio VXLAN, que incluiría uno o más VNI, para el transporte de tráfico de inquilinos sobre una infraestructura subyacente de IP común.
    - **Punto final de virtualización de red (NVE):** Dentro de un entorno NVO, el NVE se traduce como el nodo de borde del proveedor, el cual es responsable de la encapsulación/desencapsulación del tráfico de inquilinos en la red superpuesta. Y en referencia a un plano de datos VXLAN, el NVE vendría a ser el punto final del túnel virtual (VTEP)
    - **Instancia de EVPN (EVI):** Esta es una instancia de enrutamiento y reenvió dentro del dominio de EVPN que abarca e interconecta varios VTEP para proporcionar conectividad de capa 2 y capa 3 de inquilino.
    - **MAC-VRF:** una tabla de enrutamiento y reenvío virtual para almacenar direcciones de control de acceso a medios (MAC) en un VTEP para un inquilino específico.

1.5.1.1. Servicios VLAN para EVPN

Dentro de una instancia EVI se puede albergar varios dominios de difusión de capa 2 (VLAN), dichos dominios se identifican mediante una etiqueta ethernet para luego codificarla dentro de NLRI EVPN para su posterior intercambio entre los conmutadores para el aprendizaje del plano de control. Para el mecanismo de asociación de un VLAN-ID a una instancia EVI existen tres tipos de servicios VLAN definidos para EVPN: VLAN based, VLAN Bundle, y VLAN aware bundle (Chandra, 2017, p.236).

- **VLAN Based Service Interface:** Dentro de este servicio un EVI posee un solo dominio de difusión, es decir una VLAN. Por lo que se maneja un mapeo 1:1 entre la VLAN-ID y un MAC-VRF (EVI). Al hacer uso de VXLAN como mecanismo de encapsulación del plano de datos para la EVPN, se utiliza el VNI para la asignación directa al EVI dentro del servicio basado en VLAN (VLAN Based). Como ventajas en este servicio se logra un control granular y un ámbito de inundación eficiente; sin embargo, el servicio trae consigo limitaciones en el escalado (Chandra, 2017, pp.237-238). La figura 13-1, representa el mapeo 1:1 entre la VLAN-ID y la instancia EVI (VNI).

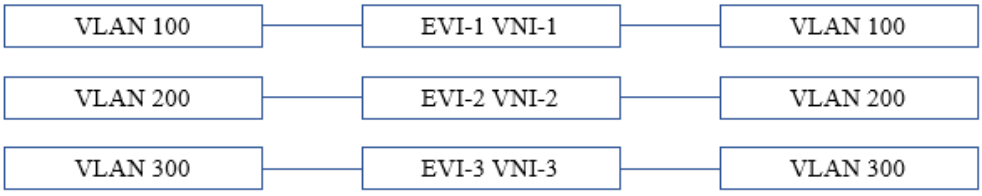

Figura 13-1: Interfaz de servicio basada en VLAN, mapeo 1:1 VLAN-VNI.

Realizado por: Mena Duval, 2021

- **VLAN Bundle Service Interface:** Dentro de este servicio un EVI adopta múltiples dominios de difusión, es decir múltiples VLANs, y a su vez todas las VLANs comparten la misma tabla de puentes para un EVI especifico. De manera que, se maneja un mapeo de N:1 entre varias VLANs y un solo MAC-VRF (EVI). Al hacer uso de VXLAN como mecanismo de encapsulación del plano de datos para la EVPN, todos y cada uno de los VLAN-ID comparten el mismo VNI que se asigna al EVI. Entre las ventajas de este servicio se encuentra una mayor facilidad de aprovisionamiento y baja sobrecarga del plano de control; sin embargo, no se tiene un control sobre el dominio de difusión del cliente, además se producen inundaciones (flooding) ineficientes ya que el tráfico BUM se inunda a todos los conmutadores que forman parte del EVI y varios de estos conmutadores pueden no tener interés ya que no tienen los VLAN-IDs previstos, lo que se traduce a que tendrán que realizar un proceso de descarte de tráfico innecesario que consumirá parte de los recursos (Chandra, 2017, 237-238). La figura 14-1, representa el mapeo N:1 entre las VLAN-IDs y la instancia EVI (VNI).

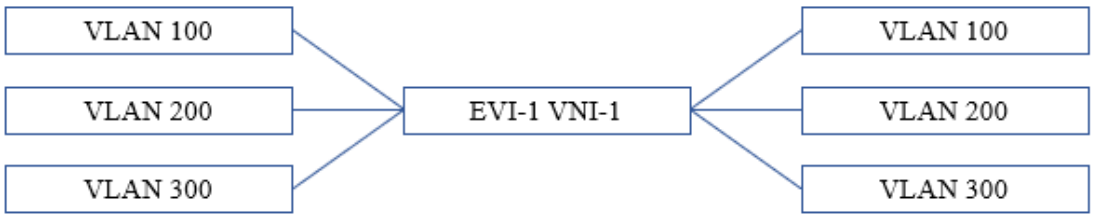

Figura 14-1: Interfaz de servicio del paquete VLAN, mapeo N:1 VLANs - VNI.

Realizado por: Mena Duval, 2021

- **VLAN Aware Bundle Service Interface:** Dentro de este servicio un EVI adopta múltiples dominios de difusión, es decir múltiples VLANs, aquí cada VLAN tiene su propia tabla de puentes que son administradas por una única MAC-VRF (EVI). Es así que existe un mapeo N:1 entre las VLAN-IDs y la MAC-VRF (EVI). Al hacer uso de VXLAN como mecanismo de encapsulación del plano de datos para la EVPN, dentro del modo consciente de la VLAN, varios VNIs se asignan al mismo EVI y las VLANs pueden agruparse con un VNI en común. Entre las ventajas de este servicio tenemos un mayor control sobre el dominio de difusión del cliente, un proceso de inundación eficiente, mayor escalado y en cuanto a desventaja una sobrecarga de aprovisionamiento (Chandra, 2017, 237-238). La figura 15-1, representa el mapeo N:1 entre las VLAN-IDs y la instancia EVI (VNI).

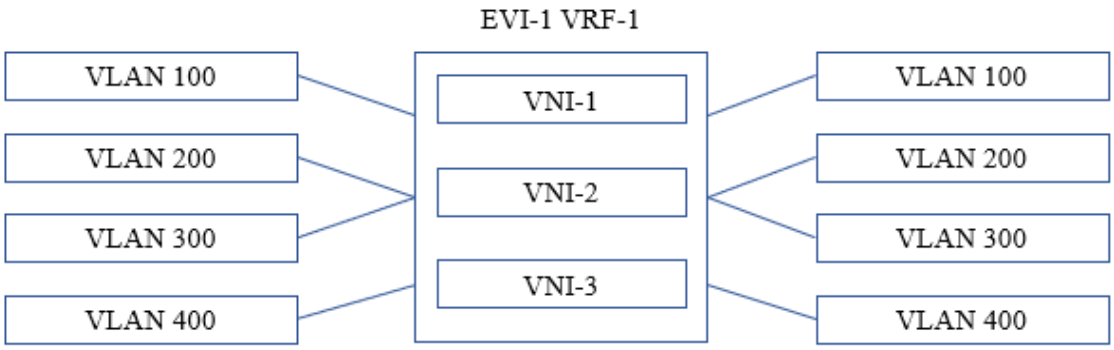

Figura 15-1: Interfaz de servicio consciente de la VLAN, mapeo N:1 VLANs - VRF.

Realizado por: Mena Duval, 2021

1.5.2. MP-BGP EVPN

La solución EVPN/VXLAN se basa en el Multiprotocolo BGP (MP-BGP). La característica principal de MP-BGP es proveer una multiplexación de familias de direcciones y NLRI relacionado en una sola sesión de peering BGP. Gracias a esto, MP-BGP permite servicios VPN a través de esta única sesión de emparejamiento BGP y proporciona lógica incorporada para separar la información de accesibilidad para los inquilinos separados (Dutt, 2018).

Dutt (2018), expresa que para que un protocolo de control intercambie información sobre redes virtuales, debe admitir tres construcciones principales:

- Una forma de identificar la dirección de red que se intercambia (el papel de AFI / SAFI)
- Identificar a qué red virtual pertenece una dirección (el rol de RD y RT)
- Identificar qué método de encapsulación de túnel se usa para construir la superposición de red virtual (indicado a través de la Comunidad Extendida de Encapsulación BGP que se usa con todos los anuncios de EVPN)

La lista AFI / SAFI que es de interés para un orador BGP se anunciará utilizando las capacidades BGP en el mensaje BGP OPEN. Dos pares de BGP intercambiarán información sobre una dirección de red solo si ambas partes anuncian un interés en su AFI / SAFI. EVPN está diseñado como una familia SAFI de L2VPN AFI. Por lo tanto, en la jerga de BGP, AFI / SAFI (25/70) de EVPN es l2vpn evpn, la tabla 5-1 muestra los valores AFI/SAFI comúnmente usados en BGP (Dutt, 2018).

Dentro del mensaje de actualización de BGP se incluyen dos atributos de comunidades extendidas (Extended communities). El primero, el atributo Route-Target se utiliza para la política de exportación/importación de rutas por parte de los conmutadores VTEP. El segundo, Encapsulation-Type define la encapsulación utilizada en el plano de datos (Tipo 8 = VXLAN) (Pasanen, 2019, p.108).

1.5.2.1. RD & RT

Route Distinguishers (RD): El RD tiene una longitud de 8 bytes y posee la siguiente codificación:

- Campo Type: 2 bytes
- Campo Value: 6 bytes

El campo tipo (Type) posee tres formatos, y según su asignación va a diferir el campo valor (Value); a su vez este último va a determinar la forma en cómo se construirá el route distinguisher. En la figura 16-1, se puede apreciar los tres tipos de formatos para construir el RD. El formato type 0: se compone del número de sistema autónomo (2 bytes) añadido un numero individual de 4 bytes de esta forma completa el RD. El formato type 1 y type 2, se asemejan en el campo valor (value) ya que tiene una longitud de 4 bytes, sin embargo, en el type 1 se asigna una dirección IP, mientras que en el type 2 se hace uso del sistema autónomo (ASN), pero al final ambos terminan precedidos por un campo numérico individual de 2 bytes (Krattiger et al., 2017). Cuando se intercambian direcciones VPN, BGP añade un RD de 8 bytes a cada dirección. Esta combinación

de RD + dirección logra hacer que la dirección sea globalmente única (Dutt, 2018). En el RFC 7432 se recomienda usar el formato Type 1 para la construcción de RD cuando se tiene una instancia MAC-VRF.

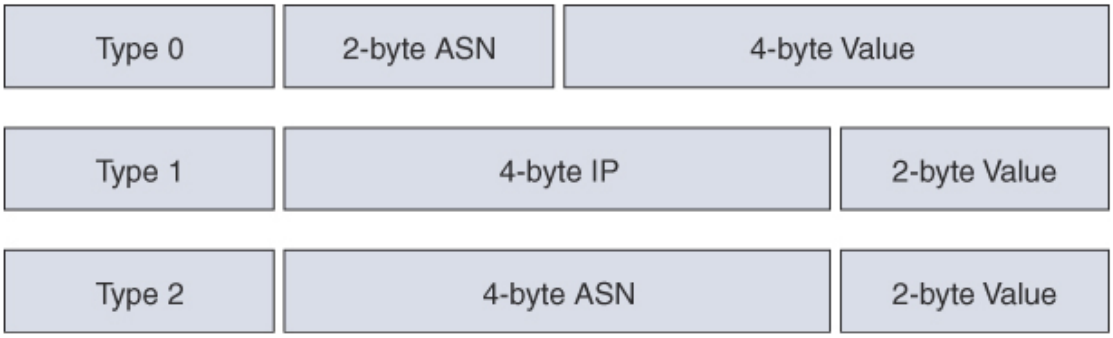

Figura 16-1: Formato para la construcción del Route Distinguisher (RD)

Fuente: (Krattiger et al., 2017)

Route Target (RT): El objetivo de la ruta (RT) es un atributo extendido en las actualizaciones de ruta de EVPN que se utiliza para controlar la distribución de la ruta en una red multi-tenant. Los VTEP de EVPN tienen una configuración de RT de importación y otra de exportación para cada VRF y cada L2VNI. Cuando un VTEP anuncia rutas EVPN, coloca su RT de exportación en la actualización de ruta. Las rutas serán recibidas por otros VTEPs en la red. Estos dispositivos compararán el valor de RT que lleva la ruta con su propia configuración local de RT de importación. Si los dos valores coinciden, la ruta será aceptada y programada en la tabla de enrutamiento. En caso contrario, la ruta no se importará (Buresh et al., 2017, p.53). Dentro del RFC 8365 se encuentra la codificación de RT para EVPN sobre VXLAN, a continuación, la figura 17-1 muestra el formato de route target (RT) para EVPN con VXLAN.

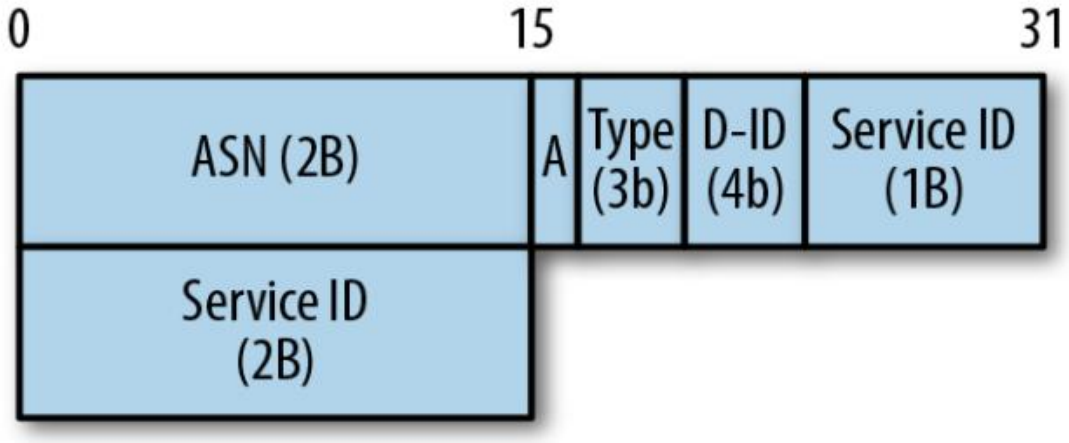

Figura 17-1: Estructura de RT para EVPN con VXLAN

Fuente: (Dutt, 2018)

El RT es un identificador de 8 bytes de longitud, comparado con RD el formato de RT se puede usar con bastante libertad, en una notación prefijo: sufijo. A continuación Dutt (2018), explica los campos contenidos en la estructura del identificador de ruta (RT):

- **ASN:** El número de sistema autónomo (ASN) de dos bytes del altavoz BGP que anuncia la dirección.
- **A:** Un bit que indica si el RT se deriva automáticamente o se configura manualmente.
- **Type:** Un campo de tres bits que indica la encapsulación utilizada en EVPN. Para VXLAN, es 1 y para VLAN, es 0.
- **ID de dominio (D-ID):** Cuatro bits que normalmente son cero. En ciertos casos, si hay una superposición en el espacio de numeración de VXLAN, este campo se utiliza para calificar el dominio administrativo al que pertenece el VNI.
- **ID de servicio:** Tres bytes que contienen el identificador de red virtual. Para VXLAN, es el VNI de tres bytes, para VLAN es de 12 bits.

1.5.2.2. Tipos de rutas BGP EVPN.

En BGP, la información de accesibilidad que esta codificada en NLRI se transporta dentro de los mensajes UPDATE. En la mayoría de combinaciones AFI/SAFI, la estructura y contenido de información de accesibilidad que son transportados por el mensaje UPDATE es la misma; sin embargo, para EVPN no es así, ya que hay distintas piezas de información que intercambiar. Por ejemplo, la actualización puede ser accesible a una dirección MAC específica o puede ser accesible a una red virtual completa, y a diferencia de IPv4 e IPv6, debido a que EVPN ya ha consumido tanto un AFI como un SAFI, no hay forma de separar la información sobre direcciones unidifusión y multidifusión. Para solventar estos problemas EVPN NLRI ofrece una clasificación por tipos de ruta. En la tabla 7-1, se muestra los diferentes tipos de rutas que establece MP-BGP EVPN (Dutt, 2018).

Tabla 7-1: Tipos de ruta establecidas dentro de EVPN.

Route Type	Lo que transporta	Funcionalidad
RT-1	Descubrimiento automático del segmento Ethernet	Admisión de múltiples terminales de host en el centro de datos, que son utilizados en reemplazo a MLAG.

RT-2	MAC, VNI, IP	Anuncia la accesibilidad a una dirección MAC específica en una red virtual y su dirección IP
RT-3	Asociación VNI/VTEP	Anuncia el interés de un VTEP en las redes virtuales.
RT-4	Transportista designado (Designated Forwarder)	Garantiza que solo un VTEP reenvíe tramas de destino múltiple a puntos finales de host múltiple
RT-5	Prefijo IP, VRF	Realiza la publicidad de prefijos IP, como rutas resumidas, y el VRF asociado al prefijo

Fuente: (Dutt, 2019, p.345)

Realizado por: Mena Duval, 2021

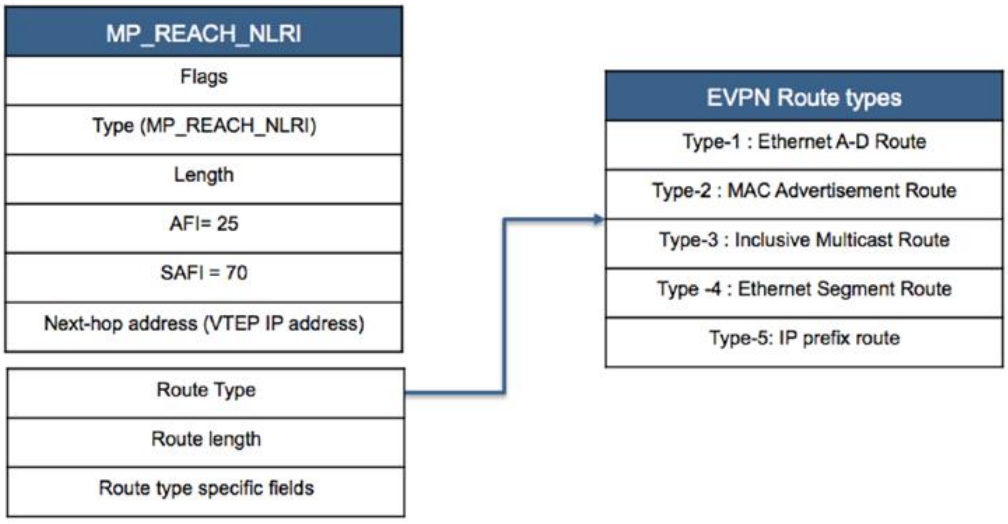

Figura 18-1: Formato de ruta EVPN NLRI

Fuente: (Arista Networks, 2021)

En la figura 18-1, se muestra el formato del atributo MP_REACH_NLRI / MP_UNREACH_NLRI, que contiene el nuevo EVPN NLRI, donde el campo Next-hop address dentro del NLRI corresponde a la dirección IP del VTEP que anuncia la ruta EVPN. Son cinco los tipos de ruta en EVPN, sin embargo las rutas que son de suma importancia para la implementación dentro de una red son RT-2, RT-3, RT-5, las demás rutas son opcionales y normalmente dentro de los conmutadores no se encuentran funcionales, a continuación el grupo de trabajo de Arista Networks (2021) describe las 3 rutas esenciales:

- **RT-2: Host advertisement route. –** El tipo de ruta 2 se usa para anunciar la accesibilidad de una dirección MAC, o un enlace de MAC/IP según lo que haya aprendido un EVI especifico.

Con esta anunciación IP, EVPN proporciona capacidad a los VTEP para que realicen la técnica de supresión ARP (ARP Suppresion, véase apartado 1.5.3.1.), para reducir el tráfico BUM dentro de la capa 2.

- **RT-3: Inclusive multicast route. –** Esta ruta se usa con el objetivo de anunciar la membresía de un dominio de capa 2 especifico, es decir, un VNI del dominio VXLAN, esto a su vez permite el descubrimiento de VTEPs remotos en un VNI específico y la población de una lista de flujo de entrada de VTEP para el reenvío del tráfico BUM.
- **RT-5: IP-prefix route advertisement. –** La ruta de tipo 5 se utiliza para anunciar prefijos de IP en lugar de las direcciones de host MAC/IP de la ruta de tipo 2. Este anuncio de prefijos en el dominio EVPN brinda la capacidad de construir topologías VPN clásicas de capa 3, esto permite una separación limpia para los altavoces BGP EVPN evitando así cualquier procesamiento relacionado con la dirección MAC para las rutas de prefijo IP anunciadas por EVPN. Además, las comunidades extendidas del tipo de ruta 5 transportan el destino de la ruta, el tipo de encapsulación y el MAC del enrutador del VTEP del siguiente salto en la superposición (Krattiger et al., 2017).

1.5.3. Mejoras de VXLAN BGP EVPN

Dentro de la tecnología EVPN existen varias mejoras para tener una mayor eficiencia junto con VXLAN, estas se describen a continuación.

1.5.3.1 Supresión ARP (ND Suppress ARP)

A diferencia del bridging tradicional, el aprendizaje de MAC entre los VTEPs dentro de EVPN se realiza en el plano de control y no en el plano de datos, esto se logra gracias a que EVPN se apoya de MP-BGP. A medida que la red va escalando también los dominios de difusión van en aumento por ende el tráfico ARP también aumenta. EVPN implementa la funcionalidad de supresión ARP que le permite reducir el flooding para el tráfico unicast desconocido. Gran parte de los hosts finales envían solicitudes ARP gratuitas (GARP) al establecer conexión, esto permite el aprendizaje local por parte de los VTEPs en la red EVPN, para luego distribuirse todas las direcciones MAC e IP aprendidas a otros VTEPs. Los NLRI de MP-BGP (anuncios MAC/IP) se utilizan para distribuir las direcciones MAC/IP entre los VTEPs y, como tal, se anuncia la alcanzabilidad, antes del tráfico a esos destinos. EVPN aplica ARP suppression para limitar las inundaciones ARP mediante la activación de la funcionalidad Proxy-ARP. Es decir, cuando un host final inicia una solicitud ARP para otro host final remoto, ésta es interceptada por el VTEP dentro de EVPN, que realiza una búsqueda Proxy-ARP en la dirección IP solicitada. Si encuentra

una coincidencia, envía una respuesta ARP en nombre del host final remoto, sin inundar innecesariamente con solicitudes ARP a las rutas que se dirigen a otros VTEPs y evitando que se consuma el ancho de banda de la red (Chandra, 2017, pp. 17-18).

1.5.3.2. Distributed IP Anycast Gateway

Si dos puntos finales que se encuentran alojados en subredes IP diferentes quieren comunicarse entre sí, necesitan de una puerta de enlace predeterminada. Para que un punto final pueda llegar a la puerta de enlace predeterminada obligatoriamente debe atravesar por una red de capa 2. La puerta de enlace anycast está enfocada al escalado horizontal reduciendo drásticamente el estado de la red y el protocolo. Por lo tanto, la implementación de la puerta de enlace anycast distribuida en cada VTEP/hoja evita que un host final tenga que atravesar un gran dominio de capa 2 para poder llegar a la puerta de enlace predeterminada. La puerta de enlace anycast aplica el concepto "uno a la asociación más cercana". La metodología que sigue se basa en que el tráfico de datos desde el host final se enruta topológicamente al nodo más cercano en un grupo de puertas de enlace que están todas identificadas por la misma dirección IP de destino. La puerta anycast se activa en cada VTEP y gracias a esto se evita el tener que adoptar otros protocolos o enviar paquetes de saludo tradicionales hacia la red (Krattiger et al., 2017).

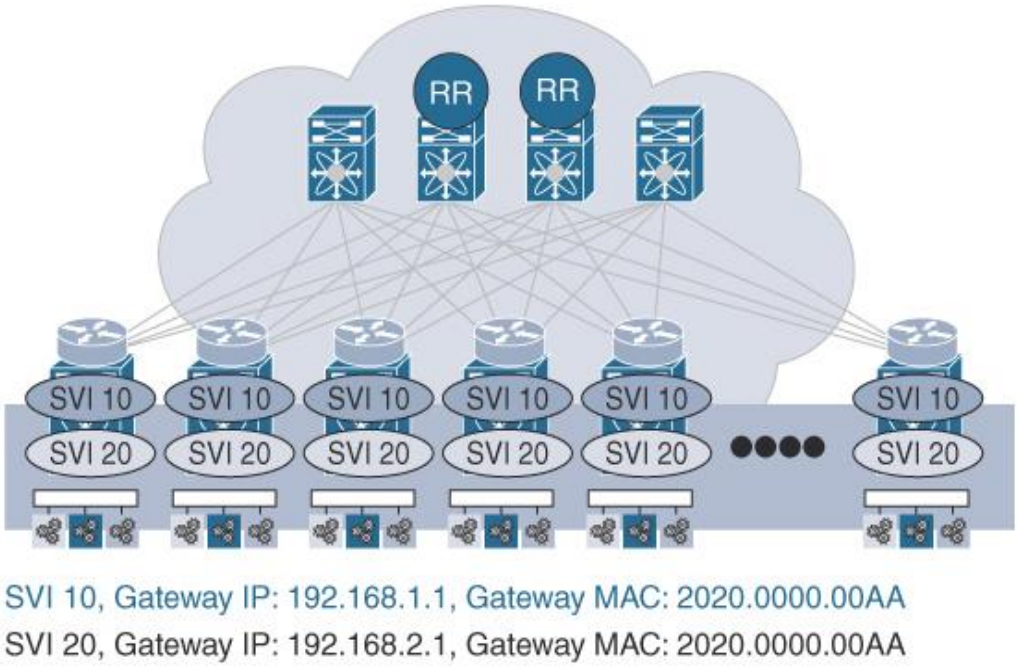

Figura 19-1: Puerta de enlace anycast distribuida.

Fuente: (Krattiger et al., 2017)

El mecanismo de escalado horizontal que adopta la puerta de enlace anycast distribuida, proporciona la puerta de enlace predeterminada más cercana a cada punto final. La dirección IP de la puerta de enlace predeterminada se comparte entre todos los dispositivos periféricos y cada dispositivo periférico es responsable de su subred IP respectiva. Sin embargo, cuando un host se mueve suele producirse un bloqueo debido al movimiento de la dirección MAC, para este problema VXLAN/EVPN implementa anycast gateway MAC addresses (AGM), donde todas las puertas de enlace anycast comparten la misma MAC en todo el tejido. De hecho, la misma AGM se comparte en todas las diferentes subredes IP, y cada subred tiene su propia IP de puerta de enlace predeterminada única. La puerta de enlace anycast no solo proporciona el enrutamiento de salida más eficiente, sino que también proporciona enrutamiento directo al VTEP donde se conecta un punto final determinado, lo que elimina la obstrucción del tráfico (Krattiger et al., 2017). La figura 19-1, muestra la realización lógica de una puerta de enlace anycast IP distribuida en una estructura VXLAN BGP EVPN.

1.5.3.3. Integrated Route and Bridge (IRB)

Dentro del borrador draft-ietf-bess-evpn-inter-subnet-forwarding publicado por la IETF se encuentran dos mecanismos que VXLAN BGP EVPN utiliza para el enrutamiento y puenteo integrado: IRB simétrico e IRB asimétrico.

IRB asimétrico sigue un patrón puente-ruta-puente (bridge–route–bridge), el cual consiste en que el tráfico que se envía desde un VTEP remoto usa un VNI diferente al tráfico de retorno del VTEP remoto. La figura 20-1, representa un ejemplo de IRB asimétrico, donde el Host A conectado al VTEP V1 desea enviar tráfico de datos hacia el Host X que se encuentra conectado al VTEP V2, ambos hosts pertenecen a subredes diferentes. Después de la resolución ARP de la puerta de enlace predeterminada el host A envía tráfico a la puerta predeterminada en la VLAN 10 y desde esta VLAN se realiza un enrutamiento hacia la VLAN 20 que esta asignado al VNI 30002, el tráfico se enruta hacia el VTEP V2 encapsulado con el VNI 30002; al llegar al VTEP V2 el tráfico se desencapsula y se puentea hacia la VLAN 20 ya que la VLAN 20 esta asignada al VNI 30002, cumpliendo así un patrón puente-ruta-puente (Krattiger et al., 2017). Debido a que los VTEP de ingreso deben configurarse con los VNI de origen y de destino, esto crea un problema de escalabilidad, ya que todos los VTEP deben configurarse con todos los VNI en la red para que puedan conocer todos los hosts conectados a esos VNI (Jain y Edgeworth, 2016, p.657).

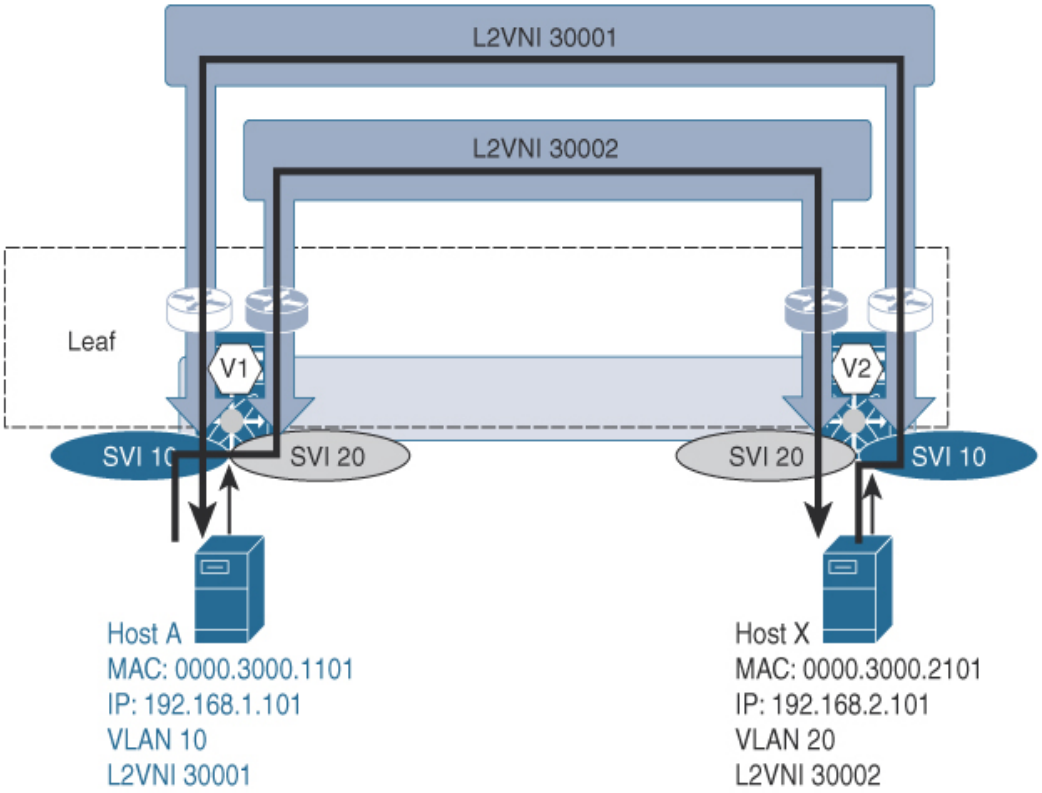

Figura 20-1: Representación del funcionamiento del IRB asimétrico.

Fuente: (Krattiger et al., 2017)

IRB simétrico es la opción más escalable, preferida y a la vez permite casos de uso adicionales que no son posibles con IRB asimétrico que brindan un gran aporte en la implementación de VXLAN BGP EVPN. IRB simétrico sigue el patrón puente-ruta-ruta-puente (bridge–route–route–bridge), puesto que todo el tráfico que entra y sale de un VTEP usa el mismo VNI (L3VNI), el cual va asociado a un VRF y que se utiliza para todo el tráfico enrutado. En la figura 21-1, se ilustra el proceso que mantiene el IRB simétrico, donde el host A conectado al VTEP V1 quiere comunicarse con el host Y perteneciente al VTEP V2, acto seguido el host A envía tráfico hacia la puerta de enlace predeterminada en la VLAN 10, y desde esa VLAN se enruta el tráfico según la IP de destino. El resultado de la búsqueda indica qué tráfico debe encapsularse en VXLAN y envía tráfico hacia VTEP V2, el cual alberga al host Y. El tráfico encapsulado con VXLAN se envía desde el VTEP V1 hacia el VTEP V2 bajo el VNI 50001, donde 50001 es el VNI de capa 3 (L3 VNI), asociado con el VRF en el que residen el Host A y el Host Y. Una vez que el tráfico VXLAN llega al VTEP V2, el tráfico se desencapsula y se enruta dentro del VRF hacia la VLAN 20 donde reside el Host Y (Krattiger et al., 2017). Cumpliéndose así la semántica de puente-ruta-ruta-puente.

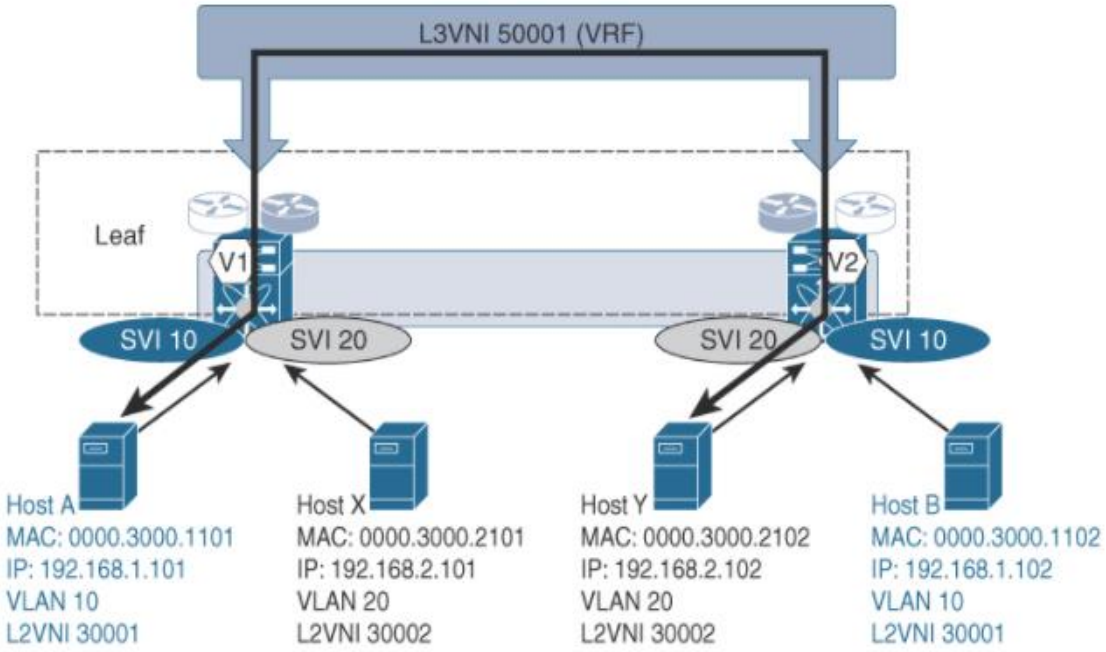

Figura 21-1: Representación del funcionamiento de IRB simétrico.

Fuente: (Krattiger et al., 2017)

1.6. Interconexión del centro de datos (DCI)

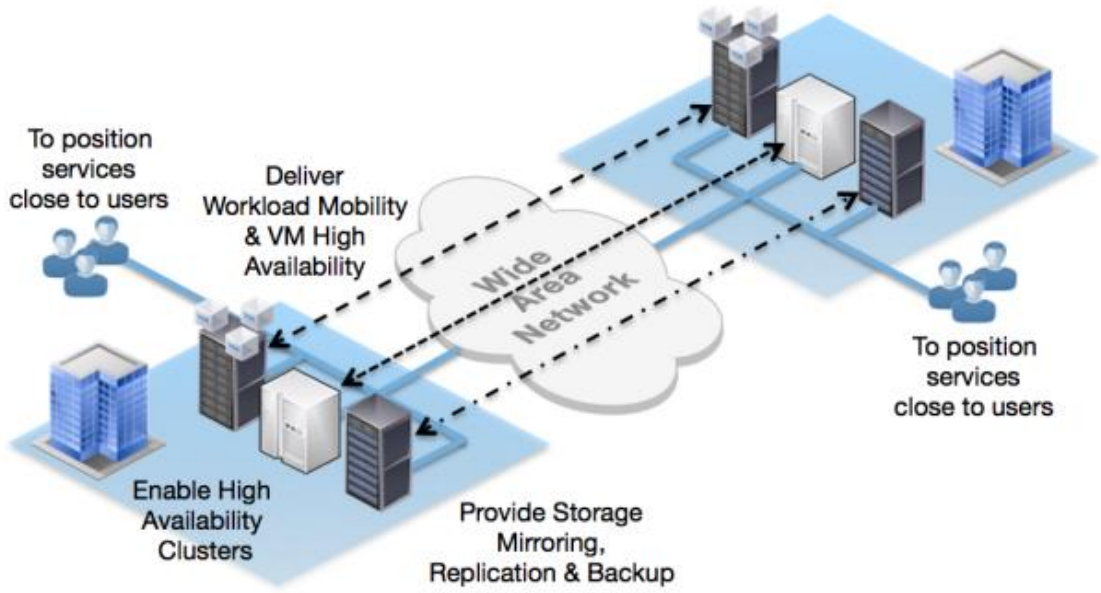

Figura 22-1: Representación de la interconexión de centros de datos.

Fuente: (Arista Networks, 2014, p.3)

La interconexión del centro de datos (DCI) es una arquitectura que permite conectar dos o más centros de datos que se encuentran distribuidos geográficamente para lograr un entorno operativo cohesivo. Gracias a DCI, la interconectividad entre centros de datos distribuidos va más allá de la simple conectividad de red, permitiendo arquitecturas de alto rendimiento y escalamiento horizontal con el objetivo de que varios centros de datos trabajen juntos al unísono de manera más eficiente. DCI permite que los centros de datos implementen sin problemas recursos de TI agrupados contenidos en la red para cumplir con los requisitos de escalabilidad. Además, DCI facilita el uso compartido de la carga de trabajo y la movilidad de los servicios, permitiendo una alta disponibilidad para que se puedan cumplir los requisitos de recuperación ante desastres, logrando así el objetivo de continuidad operativa (Pluribus Networks y Dell EMC, 2018, p.1). La figura 22-1, muestra el esquema de interconexión de centros de datos.

1.6.1. Conexión con el mundo exterior

Al hablar de la conexión con el mundo exterior de un centro de datos, que está basado en la topología de espina-hoja (Spine-Leaf), se realiza mediante el punto de interconexión externo denominado nodo de borde (border node). El nodo de borde facilita la conectividad externa tanto para el tráfico de Capa 2 como de Capa 3 a la topología de espina-hoja basada en VXLAN BGP EVPN. El nodo de borde es el equivalente a un VTEP dentro de la red VXLAN que se encarga de encapsular y desencapsular el tráfico VXLAN externo que se origina o está destinado a puntos finales debajo de los dispositivos de borde en la estructura del centro de datos. Además, el nodo de borde se encarga de reenviar el tráfico de norte a sur, lo que se traduce al tráfico desde el interior de un centro de datos hacia el mundo exterior y viceversa. Según el enfoque que maneje el centro de datos se debe tomar una decisión con respecto a la ubicación del nodo de borde, puesto que existen dos ubicaciones posibles para el nodo de borde dentro de una topología espina-hoja, estas son: espina de borde (border spine) y hoja de borde (border leaf) (Krattiger et al., 2017).

1.6.1.1. Conexión externa mediante la espina de borde (border spine)

Con esta opción de conexión externa mediante la espina de borde se obtiene un grado de eficiencia en cuanto a los flujos de trafico de norte a sur; aquí todo el tráfico de los puntos finales conectados a la hoja está a un salto de distancia cuando está destinado a llegar a las redes externas fuera de la estructura (Krattiger et al., 2017).

Se debe tomar en cuenta que en una implementación espina-hoja basada en VXLAN BGP EVPN, los conmutadores espina no están configurados para la encapsulación/desencapsulación VXLAN, sino más bien sirven como nodos de tránsito que interconectan a todas las hojas para que puedan comunicarse entre sí, por ello para utilizar la espina como espina de borde (border spine), esta

debe convertirse en un dispositivo de borde VXLAN configurado como un VTEP que encapsula y desencapsula el tráfico VXLAN para los flujos norte-sur. Sin embargo, esta espina de borde seguirá proporcionando funcionalidad subyacente al enrutar el tráfico entre los dispositivos de borde (VTEPs) según el encabezado IP externo (Krattiger et al., 2017).

Un punto importante a considerar es que todas las espinas deben estar conectadas al exterior y no solo algunas. Ya que si la conectividad con el mundo externo fuera a través de solo algunas de las espinas, esas espinas se congestionarían debido al exceso de tráfico que fluye solo a través de ellas y no por las otras espinas (Dutt, 2017). Además, esto haría que la resiliencia sea más frágil dado que perder incluso una fracción de los enlaces que se conectan a estas espinas especiales significa que esas hojas perderán el acceso completo al mundo exterior o funcionarán de manera subóptima porque su ancho de banda al mundo exterior será menor y reducido significativamente por los fallos de enlace (Dutt, 2017). La figura 23-1, muestra la forma como se conectan las espinas de borde al mundo exterior (WAN) para el tráfico norte a sur.

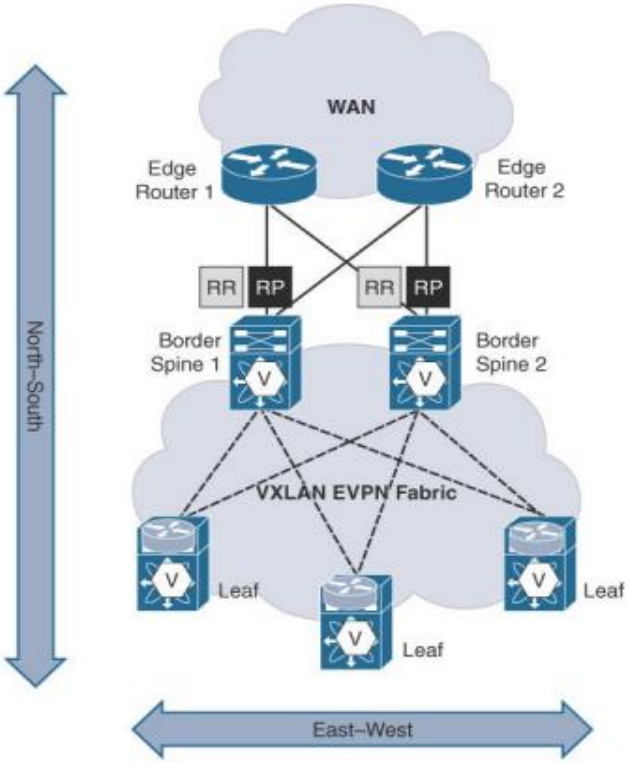

Figura 23-1: Esquema de conexión mediante espinas de borde (border spines)

Fuente: (Krattiger et al., 2017)

El modelo de conexión al mundo exterior mediante las espinas de borde, generalmente se utiliza en redes de centros de datos muy pequeños o en su defecto redes donde el tráfico de norte a sur, tiene la misma proporción del tráfico de este a oeste. Cuando el centro de datos tiene de dos a cuatro interruptores de columna (espina) y es muy sensible a los costos, es decir, no está dispuesto

a gastar dinero en interruptores de borde adicionales este es el modelo que mejor se adaptaría según las necesidades especificadas (Dutt, 2019, p.186).

1.6.1.2. Conexión externa mediante la hoja de borde (border leaf)

En esta opción de conexión externa la hoja de borde desacopla los flujos de tráfico de norte a sur y de este a oeste. En una topología Clos configurada con VXLAN BGP EVPN, las hojas de borde trabajan como puntos finales del túnel VXLAN es decir VTEPs que se encargan de encapsular y desencapsular el tráfico, por lo tanto, no requieren de consideraciones adicionales como en la conexión externa mediante las espinas. Es así que esta opción reduce el requisito de planificación de capacidad para la hoja de borde (border leaf) desde una perspectiva intra-DC, ya que la planificación solo se requiere para tráfico de norte a sur. Una consideración es que en esta opción existe un salto adicional para el tráfico de norte a sur en comparación con las espinas de borde (Krattiger et al., 2017). La principal ventaja de las hojas de borde es que aíslan el interior del centro de datos del exterior. Los protocolos de enrutamiento que se encuentran dentro del centro de datos nunca interactúan con el mundo externo, lo que proporciona una medida de estabilidad y seguridad (Dutt, 2017).

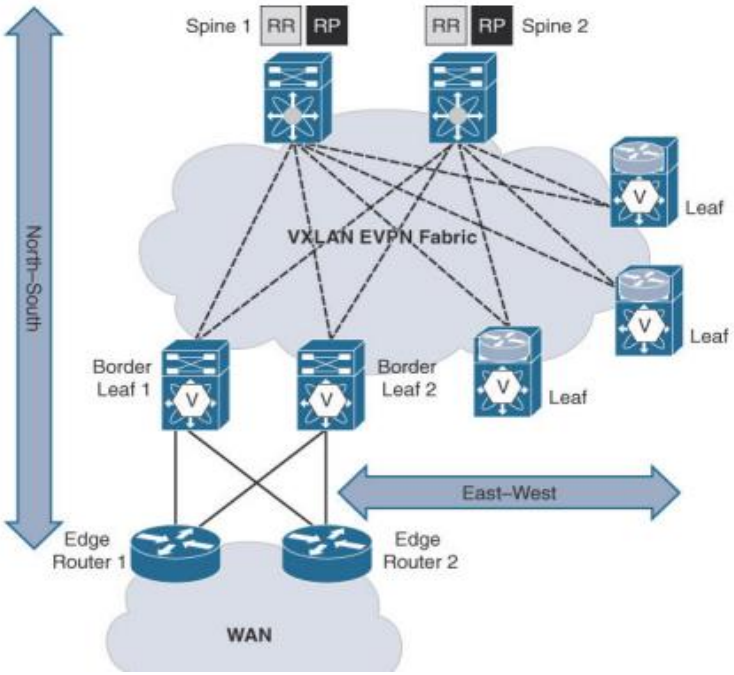

Figura 24-1: Esquema de conexión mediante hojas de borde (border leaf)

Fuente: (Krattiger et al., 2017)

La figura 24-1, muestra el esquema para la conexión externa mediante las hojas de borde. Por lo general, se implementa un conjunto de hojas de borde (border leaf), que actúan como puerta de enlace de entrada y salida de la red EVPN. Estas hojas de borde podrían emparejarse con varios

sistemas, como cortafuegos, enrutadores de Internet, enrutadores WAN, circuitos en la nube, etc., sirviendo como un punto de intercambio donde puede ocurrir la redistribución de rutas dentro y fuera de la red EVPN (Varnum, 2018).

1.6.2. Parámetros de rendimiento de la red

Para asegurar la eficiencia de una red de datos, se deben establecer y analizar ciertos parámetros que permiten determinar el rendimiento y la calidad de la comunicación dentro de la red, Chacha (2019, p.19) expone dichos parámetros a continuación:

- **Ancho de banda:** se considera como el espacio suministrado para la transmisión de datos dentro de un canal de comunicación, por lo tanto, cuanto mayor sea este espacio más cantidad de información se podrá enviar, sin embargo, aumentar el ancho de banda representan gastos elevados, además se debe considerar cambios tecnológicos necesarios para lograr el objetivo de aumentar el ancho de banda. Por tal motivo se debe realizar una administración adecuada del ancho de banda garantizando prioridades para los datos más sensibles.
- **Retardo (Delay):** es el tiempo que le toma a un paquete IP viajar desde su origen hasta un destino específico por medio del canal de comunicación. Para su cálculo intervienen parámetros como el número de nodos que tiene que atravesar el paquete, el protocolo de enrutamiento, el tráfico presente en la red, etc.
- **Jitter:** A las fluctuaciones producidas por el retraso se las conoce como jitter. El jitter es producido cuando los paquetes llegan al destino a velocidades y orden diferentes a las que se emitieron desde el origen (Ariganello y Barrientos, 2010, p.794). La variación del retardo se debe principalmente a la variación del tiempo de almacenamiento de datos en las colas (Perez, 2011, p.34).
- **Pérdida de paquetes:** es el porcentaje de paquetes que se descartan dentro del proceso de comunicación, con relación al número total de paquetes transmitidos, este efecto se presenta por el alto grado de congestión en las colas de los nodos.

1.7. Estado del Arte de BGP EVPN/VxLAN

El entorno global en función de los centros de datos ha venido acarreando una serie de hechos y sucesos que han afectado de forma negativa la normalidad de los procesos que conforman el mecanismo enfocado a la fluides de la información, siguiendo este contexto en los inicios históricos de los centros de datos se había planteado una arquitectura de tres niveles que contemplaba una supuesta mirada hacia el futuro, es decir, que era eficiente, escalable y fiable;

sin embargo, el crecimiento masivo de los datos junto con los requisitos de disponibilidad y seguridad incluyeron nuevos retos a las organizaciones que gestionan los centros de datos, es así, que a los diseños tradicionales se les ha complicado hacer frente a los numerosos retos que continuamente van surgiendo. Por tales motivos las organizaciones ya se han planteado el abandono de los tejidos Ethernet heredados iniciando una migración paulatina a tejidos IP basados en Leaf y Spine (Takamäki, 2018).

Con el avanzado mundo tecnológico virtualizado los centros de datos obtienen oportunidades de crecimiento que permiten expandirse dando cabida a nuevas sucursales para acoger a un mayor número de dispositivos conectados, creando así una necesidad de mantener sus organizaciones interconectadas, del mismo modo varias tecnologías orientadas a la interconexión fueron surgiendo entre las que destacaban las redes privadas virtuales (VPN) IP/MPLS, en las que la transparencia de la capa 2 se consigue principalmente mediante tecnologías como VPLS (Singh et al., 2017).

Sin embargo, aunque VPLS cumplía con la función de interconectar de manera satisfactoria a los centros de datos, mantenía el incomodo mecanismo de inundación y aprendizaje que durante varios años a sido el factor que aqueja a los centros de datos, para lo cual los investigadores han planteado la necesidad de buscar y evaluar nuevas tecnologías que permitan un reenvió optimo, así como una tenencia múltiple. Como solución la tecnología ethernet VPN (EVPN), ha mostrado un nivel de madurez que se adapta a las exigencias de la interconexión de los centros de datos (Makowski et al., 2019).

De esta forma Ethernet Virtual Private Network (EVPN) se propone como una tecnología emergente que aborda los desafíos de red que presentan los centros de datos geo distribuidos. Una de las principales ventajas de EVPN sobre las soluciones VPN de capa 2 heredadas es que proporciona un modo de operación totalmente activo (AA) para que el tráfico pueda ser verdaderamente multi-homed, EVPN se apoya del Multi protocolo BGP (MP-BGP) ya que incluye dentro de su familia de direcciones la integración y compatibilidad para EVPN, logrando un plano de control eficiente, que permite el anuncio de direcciones MAC, enlaces MAC e IP, así como la correcta supresión del tráfico ARP innecesario (Noghani y Kassler, 2019). EVPN tiene a su disposición varios planos de datos, no obstante, con el que mejor se adapta y que tienen como objetivo común superar las barreras impuestas en los centros de datos es la tecnología VXLAN.

La red superpuesta formada mediante el protocolo de túnel de red de área local extensible virtual (VxLAN) tiene un concepto de múltiples inquilinos, con escalabilidad de hasta 16,7 millones de segmentos de ID de VLAN, y la integración de la red privada virtual Ethernet (EVPN) que se utiliza para transportar la capa 2 e información de la capa 3 simultáneamente. De este modo mediante la combinación del tridente BGP EVPN/VxLAN se obtiene un mecanismo que logra

garantizar la seguridad de los datos en el proceso de aislamiento de la red, permitiendo maximizar la arquitectura de la red física mediante la creación de una red virtual, donde el escalado y crecimiento de los centros de datos ya no son el punto de flaqueza gracias al mecanismo de virtualización y superposición de EVPN/VxLAN (Naufal, 2021).

Es necesario señalar que las implementaciones con BGP EVPN/VxLAN en función del año 2015, tenían una inclinación a desempeñarse junto con las redes definidas por software (SDN), ya que de cierto modo la comunidad de investigadores y organizaciones afines con las telecomunicaciones, veían el gran potencial que las redes SDN podían brindar, aunque en el transcurso de los años siguientes tuvieron una decaída debido a la inmadurez, falta de implementaciones y pulido, varias implementaciones optaron por el mecanismo de EVPN/VxLAN en solitario, sin embargo, no se descarta que en un futuro próximo la forma de trabajo entre EVPN/VxLAN y SDN sea de forma nativa ya que las grandes organizaciones internacionales y fabricantes en la actualidad promueven más que nunca el uso de controladores SDN en función de las virtudes que traen consigo.

Al enfrascarse en el mecanismo BGP EVPN/VxLAN existe un punto que se debe tomar en cuenta previo a la implementación, que es la sobrecarga que trae consigo VxLAN de modo que es necesario el uso de marcos gigantes (jumbo frames), ya que se debe cumplir con uno de los requisitos principales que es evitar la fragmentación de paquetes dentro de la red con VxLAN.

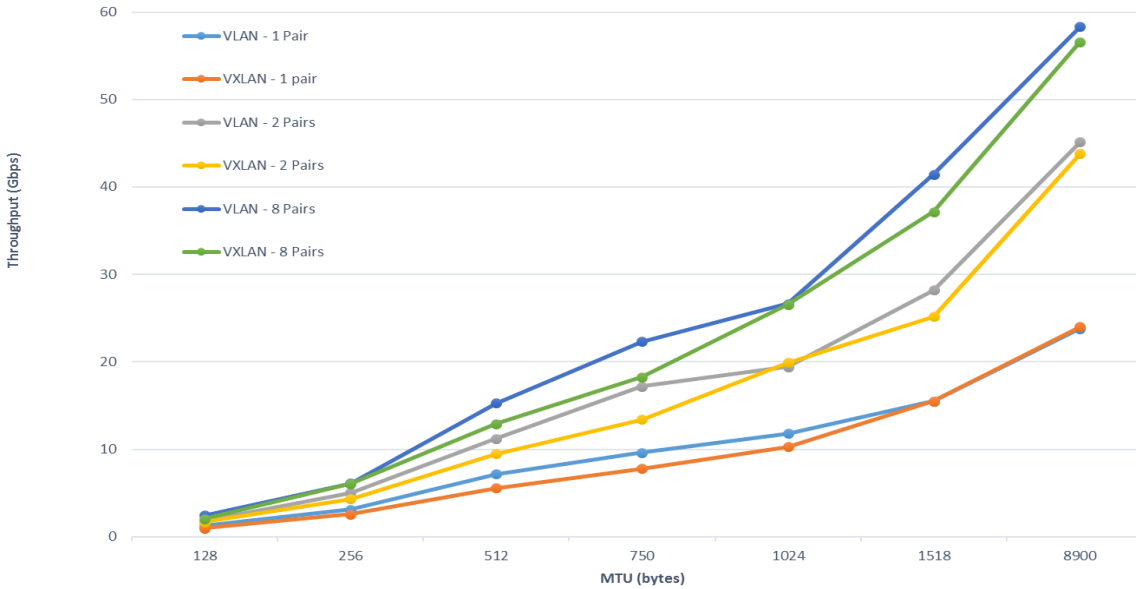

Figura 25-1: Rendimiento de VLAN frente a VxLAN antes de utilizar NIC's de próxima generación

Fuente: (Gómez, 2017)

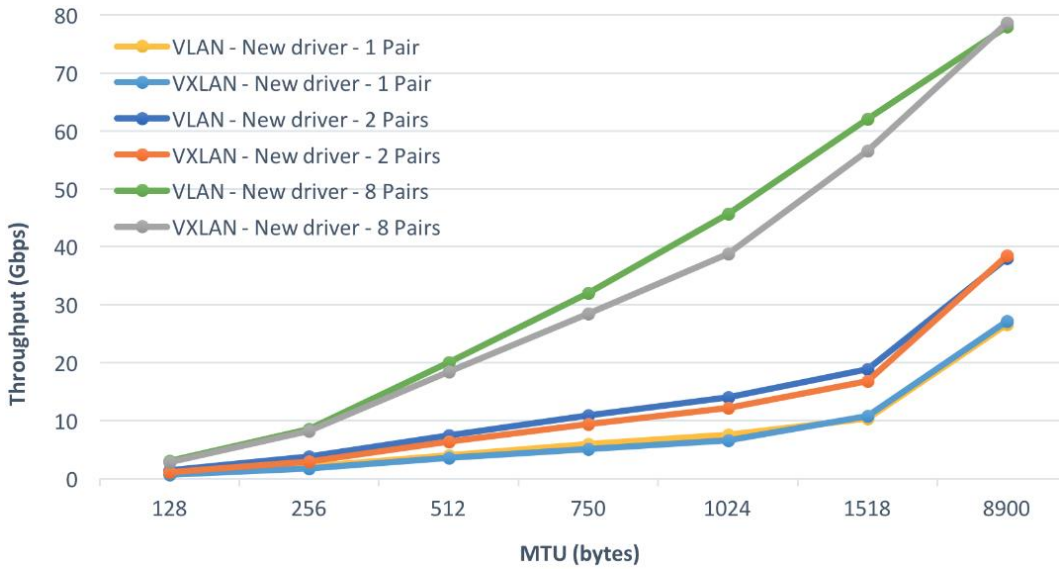

Figura 26-1: Rendimiento de VLAN frente a VxLAN después de utilizar NIC's de próxima generación

Fuente: (Gómez, 2017)

El uso de jumbo frames en las implementaciones de EVPN/VxLAN mejoran notablemente el rendimiento que se ajusta de forma positiva a los requerimientos que exigen los centros de datos para prestar servicios de forma eficiente. Como se observa en la figura 25-1, el uso nativo de Vlans presenta un mejor desempeño cuando se utilizan tarjetas de red (NIC) comunes que no están familiarizados con el Overhead que impone VxLAN, en cambio como se puede observar en la figura 26-1, al utilizar NIC's de próxima generación se logra un rendimiento a la par entre el tráfico nativo frente al superpuesto. Es así que gracias a las descargas compatibles con VXLAN disponibles en las NIC de próxima generación, se logra gozar de los beneficios de todas las ventajas derivadas de las redes superpuestas (flexibilidad, escalabilidad, neutralidad del proveedor de la red, adyacencia L2 virtual sobre cualquier topología IP, etc.) con una insignificante compensación de rendimiento (Gómez, 2017).

A nivel global EVPN/VxLAN ya está siendo implementado y conforme al uso van surgiendo mejoras y componentes que permiten obtener una mejor adaptación de la tecnología dentro y fuera de los centros de datos.

En el Ecuador hay carencias de estudios que faciliten y promuevan la adopción de la tecnología EVPN/VxLAN, por ello con la información y datos recopilados en el presente trabajo, se espera que sirva como una guía base que impulse la adopción de la tecnología de virtualización y superposición en implementaciones futuras.

CAPÍTULO II

2. MARCO METODOLÓGICO

2.1. Herramientas a utilizar

2.1.1. Simulador de red Gns3

Gns3 es una de las herramientas más utilizadas para emular, configurar, probar y solucionar problemas de redes virtuales y reales. Gns3 se caracteriza por ser de código abierto y se puede utilizar de forma gratuita por toda la comunidad que está interesada en aprender temas de redes y afines. La arquitectura de Gns3 se basa en dos componentes de software: la *Gui* denominada Gns3-all-in-one y una máquina virtual (*Vmware, Virtualbox*); su desarrollador original es Jeremy Grossman (Gns3, 2021).

En Gns3 se puede emular desde una pequeña red con unos cuantos dispositivos, hasta una gran topología de red a nivel empresarial, logrando así el objetivo de proporcionar un entorno virtual lo más cercano posible a la realidad de las redes, ya que brinda el soporte de múltiples fabricantes de dispositivos de red (*routers, switchs, firewalls, servidores, entre otros*); permitiendo la interoperabilidad entre los diferentes fabricantes para que los usuarios puedan configurar protocolos de enrutamiento y probar diferentes mecanismos de comunicación que están presentes en las telecomunicaciones y redes de datos. A continuación el grupo de trabajo de Gns3 (2021) señalan algunas de las ventajas y desventajas que tiene la aplicación:

Ventajas:

- Permite simulaciones de red en tiempo real para pruebas previas a la implementación sin necesidad de hardware de red.
- Libre creación de topologías y laboratorios personalizados dentro del entorno de Gns3, sin limitación en la cantidad de dispositivos admitidos.
- Admite imágenes de cisco VIRL y de múltiples proveedores (*Arista, Juniper, Brocade, HPE, MiKroTik).*
- Soporte nativo para Linux sin la necesidad de software de virtualización adicional.

Desventajas:

- Las imágenes de cisco no son gratuitas, puesto que para obtenerlas se debe comprar una licencia e importarlas a Gns3.

- GNS3 puede verse afectado por la configuración y las limitaciones en cuanto a CPU y memoria RAM del computador en el que es emulado.

2.1.1.1. Requisitos para el funcionamiento de Gns3

En la tabla 1-2, se exponen los requisitos que se han establecido como recomendados y óptimos para que el funcionamiento de Gns3 sea eficiente.

Tabla 1-2: Requisitos recomendados y requisitos óptimos para el funcionamiento de Gns3.

Requisitos recomendados		**Requisitos óptimos**	
Ítem	**Requisito**	**Ítem**	**Requisito**
S.O.	Windows 7 (64 bits) o posterior	*S.O.*	Windows 7 (64 bits) o posterior
Procesador	4 o más núcleos lógicos: serie AMD-V / RVI o Intel VT-X / EPT	*Procesador*	Core i7 o i9 Intel CPU / R7 o R9 AMD CPU / 8 o más núcleos lógicos - Serie AMD-V / RVI o Intel VT-X / EPT
Virtualización	Se requieren extensiones de virtualización, normalmente se habilita desde la BIOS.	*Virtualización*	Se requieren extensiones de virtualización, normalmente se habilita desde la BIOS.
Memoria	16 GB de RAM	*Memoria*	32 GB de RAM
Almacenamiento	SSD con 35 GB disponibles	*Almacenamiento*	SSD con 80 GB disponibles
Adicional	La virtualización de dispositivos requiere un uso intensivo del procesador y la memoria, es decir, más es mejor.	*Adicional*	La virtualización de dispositivos requiere un uso intensivo del procesador y la memoria, es decir, más es mejor.

Fuente: (Gns3, 2021)

Realizado por: Mena Duval, 2021

A continuación, en la tabla 2-2, se muestra las especificaciones del computador utilizado en el presente trabajo de titulación, para la construcción de la topología enfocada en la interconexión de centros de datos.

Tabla 2-2: Especificaciones del computador utilizado en el trabajo de titulación.

Ítem	**Descripción**
S.O.	Windows 10 Pro (64-bits)
Procesador	Intel(R) Core (TM) i3-10100 CPU (3.60 – 4.30 GHz), 4 núcleos, 8 hilos.
Virtualización	Habilitada
Memoria	16,0 GB (15,9 GB utilizable)
Almacenamiento	SSD 250 GB

Realizado por: Mena Duval, 2021

Tomando en referencia la tabla 1-2, los requisitos del computador usado en el presente trabajo, se encuentran en la categoría de requisitos recomendados.

2.1.2. Plataformas de enrutamiento

Utilizando el entorno de Gns3 para la confección de la topología de red se ha escogido dos tipos de plataformas de enrutamiento, entre ellas: conmutadores *Arista* y enrutadores del fabricante *Cisco*, a continuación, la tabla 3-2, detalla las características de cada uno de ellos:

Tabla 3-2: Características de las plataformas utilizadas en el diseño de la topología de red.

Ítem/Fabricante	**Arista**	**Cisco**
Versión	Arista vEOS 4.25.1F	Cisco IOSv 15.6(2)T
RAM	2048 MB	512 MB
Tipo de consola	Telnet	Telnet
CPUs	2	1
Adaptador	Intel Gigabit Ethernet (e1000)	Intel Gigabit Ethernet (e1000)
Puertos	10	6
Sistema Operativo	Arista Extensible Operating System (EOS)	Cisco IOS XE

Realizado por: Mena Duval, 2021

2.2. Diseño de la red propuesta

Tomando como referencia la figura 1-2, se puede apreciar que existen tres centros de datos identificados como PODs (vainas), estos a su vez se encuentran dispersos geográficamente. Como se describió en el apartado 1.2.1., el peso de STP hizo que la topología de tres niveles para centros de datos se volviera un dolor de cabeza para los ingenieros que administraban estas redes, por lo tanto, en la figura 1-2, cada uno de los centros de datos presenta una topología *Clos* de 2 niveles denominadas comúnmente como arquitectura *Spine-Leaf;* la cual ha mostrado notables mejoras en comparación a la topología básica de tres niveles *access-agg-core*.

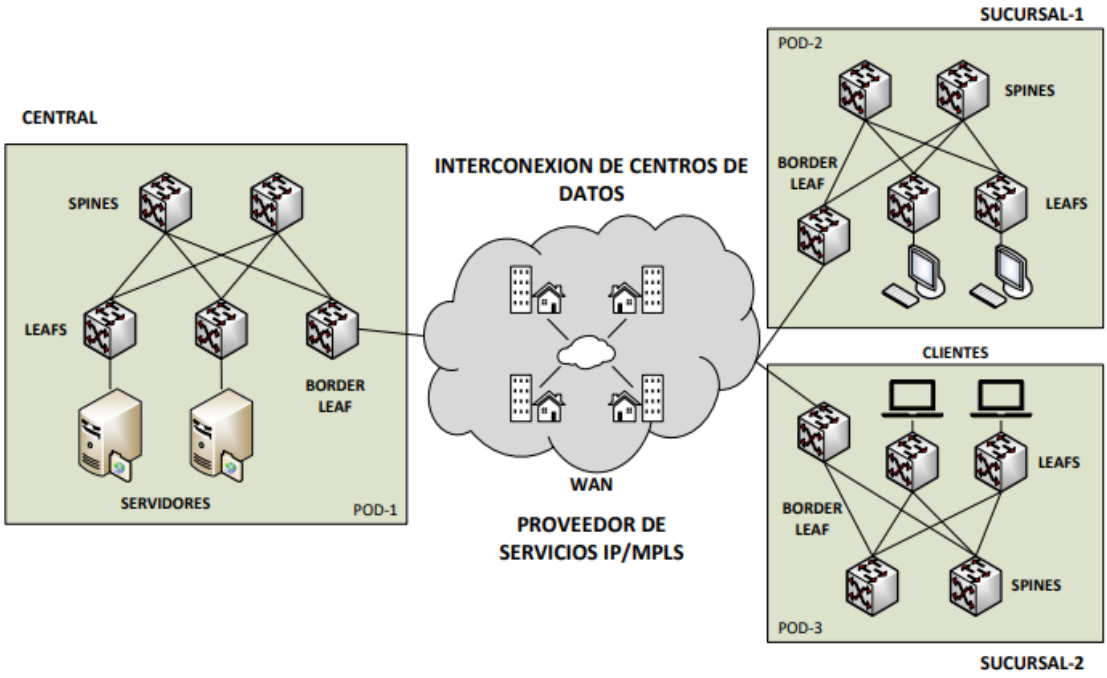

Figura 1-2: Interconexión de centros de datos basados en la topología *Spine-Leaf*

Realizado por: Mena Duval, 2021

Ahora bien, dentro de la topología *Clos*, existen dos formas posibles para lograr la interconexión de un centro de datos (DCI): interconexión mediante espinas de borde y hojas de borde. Para el desarrollo del presente trabajo de titulación se ha optado por la interconexión mediante las hojas de borde (*border-leaf*), ya que es el mecanismo mayormente usado debido a que no necesita configuraciones adicionales y tiene un impacto menor al sufrir un error y provocar una desconexión, en comparación a las espinas de borde.

Al comparar la figura 2-1 y la figura 2-2, se puede notar que las hojas de borde de la figura 2-2, representan a cada uno de los centros de datos ilustrados en la figura 2-1. Siguiendo la figura 2-2, tenemos que las hojas de borde representan a los VTEP, es decir, son el punto final del túnel

VXLAN, es así que las hojas de borde se encargan de encapsular el tráfico EVPN/VXLAN desde el servidor hacia los diferentes centros de datos que alojan a los clientes y viceversa.

Los equipos que realizarán la función de VTEP y que permitirán el funcionamiento de la superposición EVPN/VXLAN dentro de la topología son los conmutadores Arista descritos en la sección anterior; y los equipos que van a formar parte de la red WAN configurados con MPLS LDP son los enrutadores Cisco.

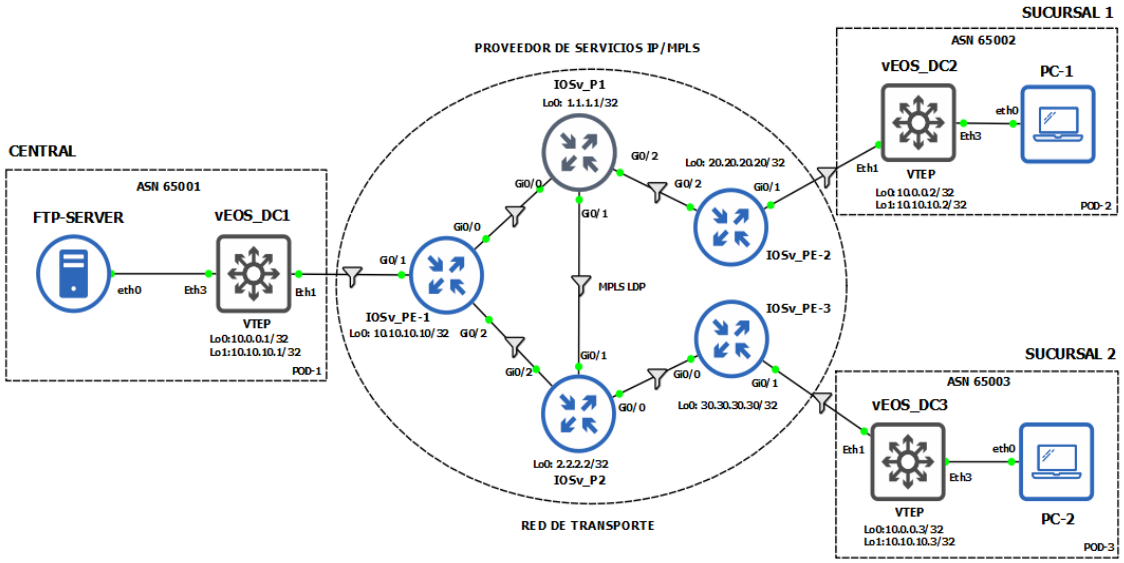

Figura 2-2: Topología de red DCI propuesta

Realizado por: Mena Duval, 2021

2.2.1. Esquema de direccionamiento

A continuación, la tabla 4-2, muestra el esquema de direccionamiento IPv4 configurado en todos y cada uno de los equipos que forman parte de la red:

Tabla 4-2: Esquema de direccionamiento para los equipos que forman parte de la red.

Equipo	*Interfaz*	*Dirección IP*	*Máscara de red*	*Gateway*	*Router-ID*	*VLAN*
Provider Core (P1)	Lo0	1.1.1.1	255.255.255.255		1.1.1.1	
	Gi0/0	10.20.2.1	255.255.255.0			
	Gi0/1	10.20.4.1	255.255.255.0			
	Gi0/2	10.20.5.1	255.255.255.0			
Provider Core (P2)	Lo0	2.2.2.2	255.255.255.255		2.2.2.2	
	Gi0/0	10.20.6.1	255.255.255.0			
	Gi0/1	10.20.4.2	255.255.255.0			
	Gi0/2	10.20.3.1	255.255.255.0			
	Lo0	10.10.10.10	255.255.255.255		10.10.10.10	

Provider Edge (PE-1)	Gi0/0	10.20.2.2	255.255.255.0			
	Gi0/1	10.10.0.1	255.255.255.252			
	Gi0/2	10.20.3.2	255.255.255.0			
Provider Edge (PE-2)	Lo0	20.20.20.20	255.255.255.255			
	Gi0/1	10.10.0.5	255.255.255.252		20.20.20.20	
	Gi0/2	10.20.5.2	255.255.255.0			
Provider Edge (PE-3)	Lo0	30.30.30.30	255.255.255.255			
	Gi0/0	10.20.6.2	255.255.255.0		30.30.30.30	
	Gi0/1	10.10.0.9	255.255.255.252			
Arista DC1-POD1 Central	Lo0	10.0.0.1	255.255.255.255			
	Lo1	10.10.10.1	255.255.255.255			
	Eth1	10.10.0.2	255.255.255.252		10.0.0.1	
	Interfaz VLAN	192.168.1.1	255.255.255.0	192.168.1.1		100
Arista DC2-POD2 Sucursal-1	Lo0	10.0.0.2	255.255.255.255			
	Lo1	10.10.10.2	255.255.255.255			
	Eth1	10.10.0.6	255.255.255.252		10.0.0.2	
	Interfaz VLAN	192.168.2.1	255.255.255.0	192.168.2.1		200
Arista DC3-POD3 Sucursal-2	Lo0	10.0.0.3	255.255.255.255			
	Lo1	10.10.10.3	255.255.255.255			
	Eth1	10.10.0.10	255.255.255.252		10.0.0.3	
	Interfaz VLAN	192.168.3.1	255.255.255.0	192.168.3.1		300
Dispositivos Finales						
Equipo	***Interfaz***	***Dirección IP***	***Anycast Gateway***	***L2VNI***	***L3VNI***	***VLAN***
FTP-SERVER	eth0	192.168.1.100/24	192.168.1.1	1000	100001	100
PC-1	eth0	192.168.2.100/24	192.168.2.1	2000	100001	200
PC-2	eth0	192.168.3.100/24	192.168.3.1	3000	100001	300

Realizado por: Mena Duval, 2021

2.2.2. Parámetros de los enlaces

Un factor importante a considerar dentro de una red, son los parámetros de los enlaces, debido a que los enlaces tienden a sufrir atenuaciones y/o degradaciones que son producidas por factores como la distancia, frecuencias, conectores, potencia entre otros, y que afectan directamente la calidad de la comunicación. Es así que Gns3 permite establecer valores relativos dentro de los enlaces para que la simulación se asemeje lo mejor posible a la realidad de las redes, esto se logra mediante la opción denominada *packet filter* la cual se ilustra en la figura 3-2. En este caso se simulan enlaces con fibra óptica, ya que son comúnmente utilizados dentro de los proveedores de servicio en el Ecuador, a continuación, se detallan dichos parámetros:

- Latencia: A una distancia cercana a 10 Km y a temperaturas de 26 °C, la fibra óptica presenta una latencia de 3 [ms] (Khatimi et al., 2019, p.8).
- Jitter: para comunicaciones mediante fibra óptica el jitter debe ser ≤ 2 [ms], de esta forma se ha establecido el valor de 2 [ms], dentro de la opción *packet filter* para cada uno de los enlaces (Northland Communications, 2013, p.3) (Jesus et al., 2015).
- Packet loss: La pérdida de paquetes en la fibra óptica suele ser una de las más bajas, por lo que está establecida como ≤ 1% de pérdida (Northland Communications, 2013, p.3) (Comcast, 2017).

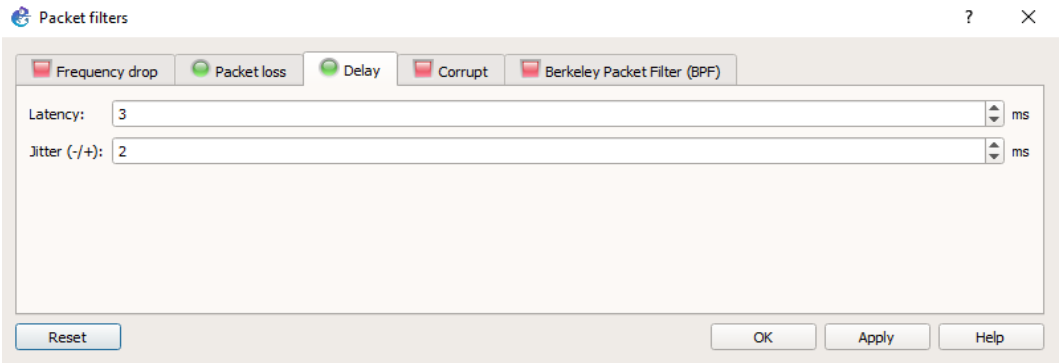

Figura 3-2: Configuración de los parámetros del enlace dentro de la opción *Packet Filter*.

Realizado por: Mena Duval, 2021

2.3. Funcionamiento de la Red DCI BGP EVPN/VXLAN

En este apartado se dividirá a la red propuesta en dos partes: en la primera parte se expone la configuración de MPLS LDP con OSPF como IGP, realizando las funciones de proveedor de servicios; en cuanto a la segunda parte se enfocará en la configuración de BGP EVPN/VXLAN con la variante de eBGP como base, permitiendo realizar el mecanismo de interconexión de los diferentes centros de datos.

2.3.1. Configuración de MPLS LDP

La figura 4-2, muestra la configuración inicial de OSPF en función del *Provider Core (P1)*, el cual funcionará como el IGP global, que permitirá el encaminamiento y alcanzabilidad entre los conmutadores Arista y enrutadores Cisco que están presentes en la red. Como se puede observar, se asignan las direcciones IP (ver Tabla 4-2) a cada una de las interfaces, a su vez estas trabajarán en el área 0 (backbone) y dentro del ID de proceso OSPF 10.

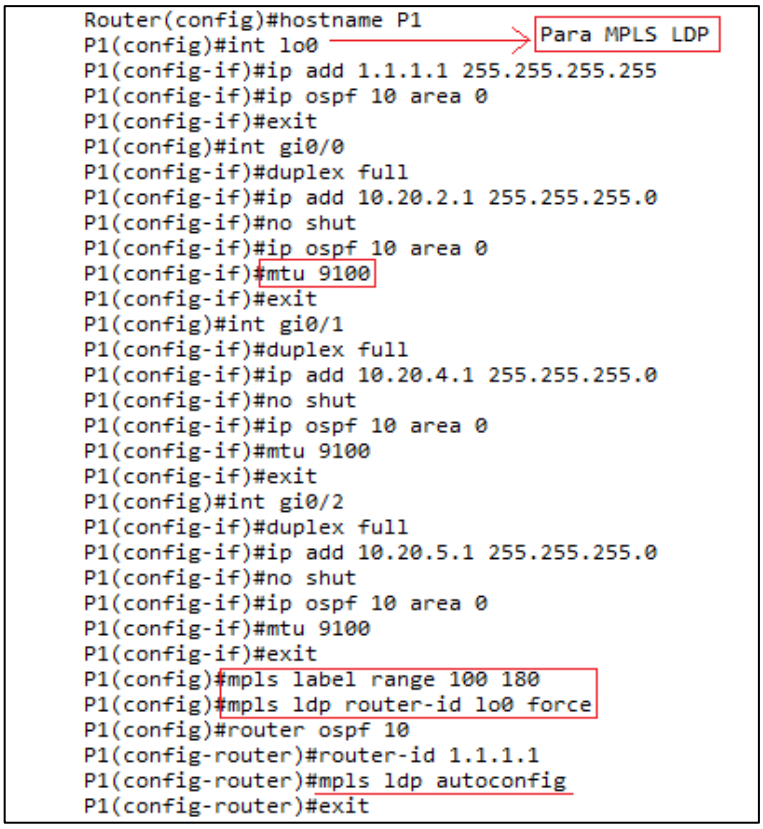

Figura 4-2: Configuración de MPLS LDP con OSPF como IGP (P1)

Realizado por: Mena Duval, 2021

Para MPLS se realiza la asignación de un rango de etiquetas; luego se hace uso de la interfaz de Loopback para que el protocolo LDP pueda enviar por ese medio sus etiquetas. MPLS sirve como red de transporte para el tráfico encapsulado que se genera mediante los centros de datos configurados con los protocolos BGP EVPN/VXLAN.

Un punto importante a considerar que esta señalado en el apartado 1.4.5., que habla sobre la sobrecarga que genera VXLAN y la importancia de hacer uso de las jumbo frames en función de los centros de datos, además del ajuste de MTU para MPLS; tomando las consideraciones respectivas se puede notar en la figura 4-2, que cada una de las interfaces están configuradas con una MTU de 9100 bytes, para que el tráfico encapsulado con EVPN/VXLAN pueda fluir sin ninguna complicación, es decir, evitando la necesidad de fragmentar los paquetes. Por último, todos los enrutadores del Proveedor de servicios (SP) llevan una configuración muy similar a los que se exponen en la figura 4-2, y difieren en pocos puntos como en las direcciones IP para cada interfaz, el Router-ID y la interfaz de loopback. La figura 5-2, muestra la tabla LFIB que confecciona MPLS para el reenvió de paquetes etiquetados con sus respectivos prefijos IP y las direcciones IP de siguiente salto (Next-hop).

```
P1#show mpls forwarding-table
Local      Outgoing   Prefix           Bytes Label   Outgoing   Next Hop
Label      Label      or Tunnel Id     Switched      interface
100        Pop Label  2.2.2.2/32       0             Gi0/1      10.20.4.2
101        Pop Label  10.20.6.0/24     0             Gi0/1      10.20.4.2
102        Pop Label  10.20.3.0/24     0             Gi0/0      10.20.2.2
           Pop Label  10.20.3.0/24     0             Gi0/1      10.20.4.2
103        Pop Label  20.20.20.20/32   0             Gi0/2      10.20.5.2
104        Pop Label  10.10.10.10/32   0             Gi0/0      10.20.2.2
105        Pop Label  10.10.0.4/30     0             Gi0/2      10.20.5.2
106        Pop Label  10.10.0.0/30     0             Gi0/0      10.20.2.2
107        205        30.30.30.30/32   0             Gi0/1      10.20.4.2
108        306        10.10.10.1/32    0             Gi0/0      10.20.2.2
109        307        10.0.0.1/32      24399         Gi0/0      10.20.2.2
110        209        10.10.0.8/30     0             Gi0/1      10.20.4.2
111        412        10.10.10.2/32    0             Gi0/2      10.20.5.2
112        413        10.0.0.2/32      23641         Gi0/2      10.20.5.2
P1#
```

Figura 5-2: Tabla LFIB correspondiente a MPLS.

Realizado por: Mena Duval, 2021

2.3.2. Configuración del tridente BGP EVPN/VXLAN

2.3.2.1. Configuración de OSPF

Como primera instancia se realiza la configuración básica de OSPF para establecer adyacencias y alcanzabilidad desde los conmutadores Arista hacia los demás enrutadores de la red, esta configuración se muestra en la figura 6-2.

```
CENTRAL(config)#ip routing
CENTRAL(config)#int eth1
CENTRAL(config-if-Et1)#no switchport
CENTRAL(config-if-Et1)#mtu 9100
CENTRAL(config-if-Et1)#ip add 10.10.0.2/30
CENTRAL(config-if-Et1)#no shut
CENTRAL(config-if-Et1)#exit
CENTRAL(config)#interface loopback 0
CENTRAL(config-if-Lo0)#ip address 10.0.0.1/32
CENTRAL(config-if-Lo0)#exit
CENTRAL(config)#interface loopback 1
CENTRAL(config-if-Lo1)#ip address 10.10.10.1/32
CENTRAL(config-if-Lo1)#exit
CENTRAL(config)#router ospf 10
CENTRAL(config-router-ospf)#router-id 10.0.0.1
CENTRAL(config-router-ospf)#log-adjacency-changes
CENTRAL(config-router-ospf)#network 10.10.0.0/30 area 0
CENTRAL(config-router-ospf)#network 10.0.0.1/32 area 0
CENTRAL(config-router-ospf)#network 10.10.10.1/32 area 0
CENTRAL(config-router-ospf)#exit
```

AFI/SAFI 1
EVPN

AFI/SAFI 1
VTEP VXLAN

Figura 6-2: Configuración de OSPF en los conmutadores Arista

Realizado por: Mena Duval, 2021

- Se asigna la respectiva dirección IPv4 a la interfaz eth1 del conmutador Arista.
- En la interfaz de salida hacia la red del proveedor de servicios se ajusta una MTU de 9100 bytes, permitiendo el tráfico mediante jumbo frames y tomando en cuenta la sobrecarga de VXLAN.

- La interfaz Loopback 0, se utilizará para el establecimiento de sesiones BGP EVPN, para mensajes de actualización, familia de direcciones AFI/SAFI IPv4; en general para el plano de control.
- La interfaz Loopback 1, se anuncia en BGP dentro de la familia de direcciones AFI/ SAFI IPv4 = 1, y se utiliza en los VTEP de VXLAN.
- Dentro del proceso OSPF 10, se anuncian las interfaces configuradas con la respectiva área de trabajo.

2.3.2.2. Configuración de parámetros básicos BGP

Una vez que se establece la alcanzabilidad entre los routers y conmutadores en la red, se procede a configurar los parámetros básicos del protocolo de borde BGP, como se muestra en la figura 7-2.

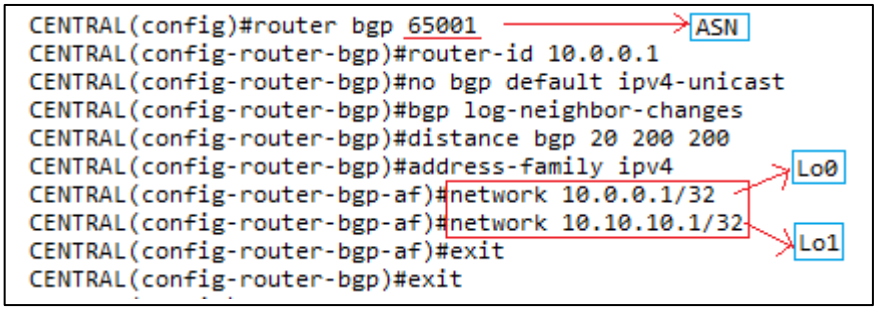

Figura 7-2: Configuración de parámetros básicos previos para BGP.

Realizado por: Mena Duval, 2021

- A cada conmutador Arista se le asigna un número de sistema autónomo puesto que BGP trabajará con su variante de eBGP dando soporte de capa subyacente (Underlay) para los protocolos EVPN/VXLAN, la interfaz loopback 0 trabaja conjuntamente con BGP para el intercambio de tráfico, es por ello que el *router-id* se establece con la interfaz *Lo0* y posteriormente es anunciada en la familia de direcciones AFI/SAFI IPv4=1.
- Se configuran comandos básicos, por ejemplo:
  - *no bgp default ipv4-unicast:* con este comando se desactiva el comportamiento predeterminado o común que tiene BGP para trabajar con la familia de direcciones *ipv4 unicast*, ya que BGP pasará de su estado simple a trabajar como Multiprotocolo BGP.
  - *bgp log-neighbor-changes:* Se habilita este comando para establecer cualquier comportamiento o cambios de estado en los vecinos, esto con el propósito de dar estabilidad a la red y tener opciones para solucionar problemas de red.

- *distance bgp 20 200 200:* el objetivo de este comando es dar prioridad siempre a eBGP sobre iBGP en función de las distancias administrativas.

- Por último, se hace el anuncio de las interfaces de loopback en la familia de direcciones IPv4 AFI/SAFI.

Una buena recomendación para mantener una red estable y eficiente, es la correcta configuración de protocolos, mecanismos y equipos; ese es el enfoque de la configuración inicial de BGP, para que los acoples entre BGP y EVPN sean precisos.

2.3.2.3. Configuración del plano de control MP-BGP EVPN

Como se muestra en la figura 8-2, se hace uso del comando *service routing protocols model multi-agent*, que es propio de los conmutadores arista y su funcionalidad es habilitar las capacidades EVPN. Una vez activado EVPN se procede a configurar la superposición MP-BGP EVPN como plano de control para VXLAN.

```
CENTRAL(config)#service routing protocols model multi-agent⇦ Habilitar Evpn
CENTRAL(config)#router bgp 65001
CENTRAL(config-router-bgp)#neighbor evpn peer group
CENTRAL(config-router-bgp)#neighbor evpn next-hop-unchanged
CENTRAL(config-router-bgp)#neighbor evpn update-source loopback 0
CENTRAL(config-router-bgp)#neighbor evpn ebgp-multihop 5
CENTRAL(config-router-bgp)#neighbor evpn send-community extended
CENTRAL(config-router-bgp)#neighbor evpn maximum-routes 12000 warning-only
CENTRAL(config-router-bgp)#neighbor 10.0.0.2 peer group evpn
CENTRAL(config-router-bgp)#neighbor 10.0.0.2 remote-as 65002
CENTRAL(config-router-bgp)#neighbor 10.0.0.3 peer group evpn
CENTRAL(config-router-bgp)#neighbor 10.0.0.3 remote-as 65003
CENTRAL(config-router-bgp)#address-family evpn
CENTRAL(config-router-bgp-af)#neighbor evpn activate
CENTRAL(config-router-bgp-af)#exit
CENTRAL(config-router-bgp)#exit
```

Figura 8-2: Configuración de la superposición MP-BGP EVPN como plano de control para VXLAN

Realizado por: Mena Duval, 2021

- Se establece una plantilla de grupo de pares (*peer-group*), en esencia este comando sirve para aplicar una lista de configuraciones a varios vecinos, reduciendo la cantidad de configuración. En este caso el *peer group* se establece con el nombre de *evpn* y lleva consigo las siguientes configuraciones:

    - *Next-hop-unchanged:* En una sesión eBGP, de forma predeterminada, el enrutador cambia el atributo de siguiente salto de una ruta BGP (a su propia dirección) cuando el enrutador envía una ruta. Si la función BGP *Next Hop Unchanged* está configurada,

BGP enviará rutas a un par de múltiples saltos eBGP sin modificar el atributo de siguiente salto (Cisco, 2019a).

- Con el comando *update-source loopback 0*, BGP establece que los pares se establecerán por medio de la interfaz de loopback y será la fuente para establecer sesiones TCP entre pares BGP.
- *ebgp-multihop 5:* Con este comando se establece el TTL=5, que representa el número de saltos entre los enrutadores para llegar a su destino dentro de la red propuesta, si el TTL=0, el paquete se descarta.
- *send-community extended:* Este comando permite activar los servicios VPN, es decir, para que los atributos se puedan compartir entre pares EVPN; este comando va ligado con el comando *route-target*.
- *Máximum-routes 12000 warning only:* en este comando se establece la cantidad de rutas BGP que acepta el conmutador de un vecino específico, en este caso el valor de 12000 es un valor predeterminado dentro de Arista y al llegar al límite se genera un mensaje de error con la acción *warning only*.

- Una vez finalizada la plantilla *evpn peer group*, se procede a asignar las configuraciones a los vecinos mediante su dirección de *loopback 0* con su respectivo sistema autónomo.
- Por último, se procede a activar a todos los vecinos que forman parte de la plantilla *evpn peer group* dentro de la familia de direcciones AFI/SAFI correspondiente a EVPN, es decir, se establece el plano de control BGP EVPN; la figura 9-2, muestra un resumen de los vecinos que se han establecido en las sesiones BGP EVPN por medio del comando *show bgp evpn summary*.

```
CENTRAL(config)#show bgp evpn summary
BGP summary information for VRF default
Router identifier 10.0.0.1, local AS number 65001
Neighbor Status Codes: m - Under maintenance
  Neighbor         V  AS           MsgRcvd   MsgSent  InQ OutQ  Up/Down State   PfxRcd PfxAcc
  10.0.0.2         4 65002              9         9    0    0 00:05:32 Estab   0      0
  10.0.0.3         4 65003              9        10    0    0 00:05:00 Estab   0      0
CENTRAL(config)#
```

Figura 9-2: Vecinos establecidos dentro de la sesión BGP EVPN

Realizado por: Mena Duval, 2021

2.3.2.4. Configuración de VLAN e instancia VRF

Antes de la configuración de los VTEPs, se procede a configurar la VLAN con su respectiva instancia de VRF, como se muestra en la figura 10-2.

```
CENTRAL(config)#vrf instance Tesis_DCI
CENTRAL(config-vrf-Tesis_DCI)#ip routing vrf Tesis_DCI
CENTRAL(config)#ip virtual-router mac-address aa:aa:00:00:00:aa
CENTRAL(config)#vlan 100
CENTRAL(config-vlan-100)#name v100
CENTRAL(config-vlan-100)#exit
CENTRAL(config)#interface vlan 100
CENTRAL(config-if-Vl100)#mtu 9100
CENTRAL(config-if-Vl100)#vrf Tesis_DCI
CENTRAL(config-if-Vl100)#ip address virtual 192.168.1.1/24
CENTRAL(config-if-Vl100)#exit
CENTRAL(config)#interface eth3
CENTRAL(config-if-Et3)#mtu 9100
CENTRAL(config-if-Et3)#switchport mode access
CENTRAL(config-if-Et3)#switchport access vlan 100
CENTRAL(config-if-Et3)#no shut
CENTRAL(config-if-Et3)#exit
CENTRAL(config)#
```

Figura 10-2: Configuración de VLAN e instancia VRF

Realizado por: Mena Duval, 2021

- Como primer punto se crea una instancia VRF con el nombre Tesis_DCI, acto seguido se habilita la funcionalidad de enrutamiento IP.
- Se crea la VLAN 100 y se le asigna un nombre en este caso *v100.*
- Dentro de la configuración de la interfaz VLAN 100 se establece el valor de MTU con el que se va a trabajar en este caso es 9100 bytes, se realiza una asociación VLAN-VRF, donde de ahora en adelante la VLAN 10 pasara a ser representada por la VRF denominada Tesis_DCI.
- Prosiguiendo mediante el comando *ip address virtual* se asigna la respectiva dirección IPv4 para la interfaz vlan 100, este a su vez cumple la función de puerta de enlace anycast distribuida (Distributed IP Anycast Gateway), descrita en la sección 1.5.3.2; y que está estrechamente relacionado con la configuración *ip virtual-router mac-address*, cuyo enfoque principal es permitir la movilidad de los puntos finales hacia otros centros de datos.
- Por último, en la interfaz *eth3*, se permite el acceso de la *vlan 100*, para que todo el tráfico pueda fluir a través de esta interfaz.

2.3.2.5. Configuración del plano de datos VXLAN

```
CENTRAL(config)#int vxlan 1
CENTRAL(config-if-Vx1)#vxlan source-interface loopback 1
CENTRAL(config-if-Vx1)#vxlan udp-port 4789
CENTRAL(config-if-Vx1)#vxlan learn-restrict any
CENTRAL(config-if-Vx1)#vxlan vlan 100 vni 1000 ---> L2VNI
CENTRAL(config-if-Vx1)#vxlan vrf Tesis_DCI vni 100001 ---> L3VNI
CENTRAL(config-if-Vx1)#exit
```

Figura 11-2: Configuración de VTEP para la encapsulación VXLAN

Realizado por: Mena Duval, 2021

- Dentro de la interfaz *vxlan 1*, se procede a configurar los diferentes parámetros con los que van a trabajar los VTEP:
    - *Vxlan source-interface loopback 1:* previamente en las secciones 2.3.2.1. y 2.3.2.2., se estableció la interfaz *Lo1* y se habilitó en la familia de direcciones AFI/SAFI IPv4 dentro de eBGP, una práctica recomendada es configurar una interfaz de bucle invertido que este destinada únicamente a los procesos que conllevan los VTEP y que no está ligada a las tareas de protocolos de capa 3, por ello la interfaz *Lo1* solo trabaja para el VTEP.
    - El puerto de destino establecido por la IANA y considerado como predeterminado es el puerto de destino 4789 para VXLAN que se encapsula en IP/UDP, sin embargo, el puerto puede ser configurado para otros propósitos según los requerimientos de las aplicaciones o host finales.
    - *Vxlan learn-restrict any:* aplicando este comando se evita que los VTEP se agreguen en listas de inundaciones dinámicas, cuando el tráfico VXLAN se recibe de fuentes no confiables (Arista Networks, 2021).
    - Para la *vlan 100* se configura el *vni* de capa 2 (*L2VNI*), que tiene mayor funcionalidad dentro del mismo segmento de capa 2, es decir, dentro de un mismo dominio de un VTEP o dentro de un mismo centro de datos; no es recomendable utilizar *L2VNI* para transportar el tráfico fuera del centro de datos. Por lo tanto, para la interconexión (DCI), se hace uso de los *L3VNI* que representan a las VRF para el enrutamiento fuera del centro de datos, una VRF puede contener varios *L2VNI*, pero una VRF es asignada a un solo *L3VNI*.

Hasta este punto ya se ha configurado tanto el plano de control como el plano de datos, el último paso es configurar el acople entre BGP EVPN/VXLAN para la forma de reenvió de información, previo a esta configuración, se realiza una acotación en referencia a como los VTEPs van a realizar el descubrimiento de otros VTEP remotos y direcciones MAC.

Los modelos actuales que se manejan a través del mecanismo de inundación y aprendizaje operan con un plano de control de multidifusión o con replicación de ingreso, donde el administrador de red configura manualmente los VTEP remotos en la lista de inundaciones. Ambos están controlados por un plano de datos, es decir, los MAC se aprenden a través de la inundación. En cambio, un aprendizaje BGP EVPN MAC es un plano de control basado en estándares (MP-BGP) que se utiliza para descubrir a los VTEPs remotos y anunciar direcciones MAC y enlaces MAC / IP en la superposición de VXLAN, de esta forma se eliminan los paradigmas de inundación y aprendizaje de los enfoques ya mencionados (multidifusión o HER). Como enfoque basado en estándares, el descubrimiento y por lo tanto la publicidad de los modelos de servicio EVPN pueden interactuar entre múltiples proveedores. Esto destaca una ventaja importante y poderosa de BGP EVPN; es decir, es un plano de control único para múltiples encapsulaciones de plano de

datos y define los servicios VPN de capa 2 y capa 3; con esto, a medida que los operadores de red avanzan hacia la simplicidad y la automatización, tener un protocolo de plano de control y una familia de direcciones para todos los planos de datos y servicios VPN resultará extremadamente poderoso (Arista Networks, 2021). En la figura 12-2, se puede observar que dentro de la interfaz *vxlan* 1, el encargado del aprendizaje MAC es EVPN, además se muestra la asignación de *L2VNI & L3VNI*, para la vlan y para la instancia VRF respectivamente.

```
CENTRAL(config)#show interface vxlan 1
Vxlan1 is up, line protocol is up (connected)
  Hardware is Vxlan
  Source interface is Loopback1 and is active with 10.10.10.1
  Replication/Flood Mode is headend with Flood List Source: EVPN
  Remote MAC learning via EVPN
  VNI mapping to VLANs
  Static VLAN to VNI mapping is
    [100, 1000]
  Dynamic VLAN to VNI mapping for 'evpn' is
    [4094, 100001]
  Note: All Dynamic VLANs used by VCS are internal VLANs.
        Use 'show vxlan vni' for details.
  Static VRF to VNI mapping is
   [Tesis_DCI, 100001]
  MLAG Shared Router MAC is 0000.0000.0000
CENTRAL(config)#
```

Figura 12-2: Configuración de la interfaz VXLAN parámetros para los VTEP.

Realizado por: Mena Duval, 2021

2.3.2.6. Configuración de RT & RD en BGP EVPN/VXLAN

En la sección denominada *"LOCAL",* se configura la forma en como VXLAN va a enviar sus paquetes encapsulados por medio del plano de control conformado por MP-BGP EVPN, dentro del sistema autónomo respectivo se configura la interfaz *vlan*, se procede a configurar el *route-distinguisher (RD→ASN: VNI)*, que se antepone a los prefijos y permite crear una dirección única a nivel mundial; el *route-target (RT→ VLAN: VNI)* se encarga de la importación y exportación de rutas, esta primera configuración representa el tipo de ruta 2 (RT-2) dentro de *EVPN* y su funcionalidad reside dentro del segmento de capa 2.

```
CENTRAL(config)#router bgp 65001
CENTRAL(config-router-bgp)#vlan 100
CENTRAL(config-macvrf-100)#rd 65001:1000
CENTRAL(config-macvrf-100)#route-target both 100:1000
CENTRAL(config-macvrf-100)#redistribute learned
CENTRAL(config-macvrf-100)#exit
CENTRAL(config-router-bgp)#vrf Tesis_DCI
CENTRAL(config-router-bgp-vrf-Tesis_DCI)#rd 10.0.0.1:1
CENTRAL(config-router-bgp-vrf-Tesis_DCI)#route-target export evpn 1:100001
CENTRAL(config-router-bgp-vrf-Tesis_DCI)#route-target import evpn 1:100001
CENTRAL(config-router-bgp-vrf-Tesis_DCI)#redistribute connected
CENTRAL(config-router-bgp-vrf-Tesis_DCI)#exit
CENTRAL(config-router-bgp)#exit
CENTRAL(config)#
```

LOCAL

DCI

Figura 13-2: Configuración de los atributos RT & RD para BGP EVPN/VXLAN

Realizado por: Mena Duval, 2021

La figura 14-2, muestra las características de la instancia EVPN (EVI), en este caso el tipo de servicio de VLAN que se está utilizando es el VLAN Based Service Interface (véase sección 1.5.1.1.), y el tipo de encapsulación es VXLAN.

```
CENTRAL(config)#show bgp evpn instance
EVPN instance: VLAN 100
  Route distinguisher: 65001:1000
  Route target import: Route-Target-AS:100:1000
  Route target export: Route-Target-AS:100:1000
  Service interface: VLAN-based
  Local IP address: 10.10.10.1
  Encapsulation type: VXLAN
CENTRAL(config)#
```

Figura 14-2: Características de la instancia EVPN (EVI).

Realizado por: Mena Duval, 2021

En la sección denominada como *"DCI"*, se realiza la configuración de la *VRF Tesis_DCI*, que es la encargada procesar todo el tráfico que se genera en la instancia EVI y que se reenviara hacia otros centros de datos, la forma en como están configurados los parámetros *RD & RT*, se considera como una configuración recomendada por los administradores de red, ya que no incorpora ningún elemento del centro de datos de origen, es este caso ni la *vlan* ni el *Asn*.

```
CENTRAL(config)#show ip route vrf Tesis_DCI

VRF: Tesis_DCI
Codes: C - connected, S - static, K - kernel,
       O - OSPF, IA - OSPF inter area, E1 - OSPF external type 1,
       E2 - OSPF external type 2, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type2, B - BGP, B I - iBGP, B E - eBGP,
       R - RIP, I L1 - IS-IS level 1, I L2 - IS-IS level 2,
       O3 - OSPFv3, A B - BGP Aggregate, A O - OSPF Summary,
       NG - Nexthop Group Static Route, V - VXLAN Control Service,
       DH - DHCP client installed default route, M - Martian,
       DP - Dynamic Policy Route, L - VRF Leaked,
       RC - Route Cache Route

                                                                  MAC VTEP's
Gateway of last resort is not set

 C        192.168.1.0/24 is directly connected, Vlan100
 B E      192.168.2.100/32 [20/0] via VTEP 10.10.10.2 VNI 100001 router-mac 0c:8b:19:b4:90:bf
 B E      192.168.2.0/24 [20/0] via VTEP 10.10.10.2 VNI 100001 router-mac 0c:8b:19:b4:90:bf
 B E      192.168.3.100/32 [20/0] via VTEP 10.10.10.3 VNI 100001 router-mac 0c:8b:19:bb:ee:55
 B E      192.168.3.0/24 [20/0] via VTEP 10.10.10.3 VNI 100001 router-mac 0c:8b:19:bb:ee:55

CENTRAL(config)#
```

Figura 15-2: Prefijos aprendidos a través de VTEPs remotos pertenecientes a la VRF denominada Tesis_DCI por medio del plano de control BGP EVPN.

Realizado por: Mena Duval, 2022

En esencia es aquí donde se establece la asociación entre el plano de control conformado por MP-BGP EVPN y el plano de datos VXLAN. La figura 16-2, muestra el resultado de la aplicación del tridente *BGP EVPN/VXLAN* para la interconexión de centros de datos en función de la topología propuesta, la cual utiliza el tipo de ruta 5 (RT-5) que es propio de EVPN por medio de una instancia VRF denominada *Tesis_DCI* asociada al *L3VNI 100001*, y la forma de enrutamiento y puenteo basado en IRB simétrico que esta descrito en el apartado 1.5.3.3.

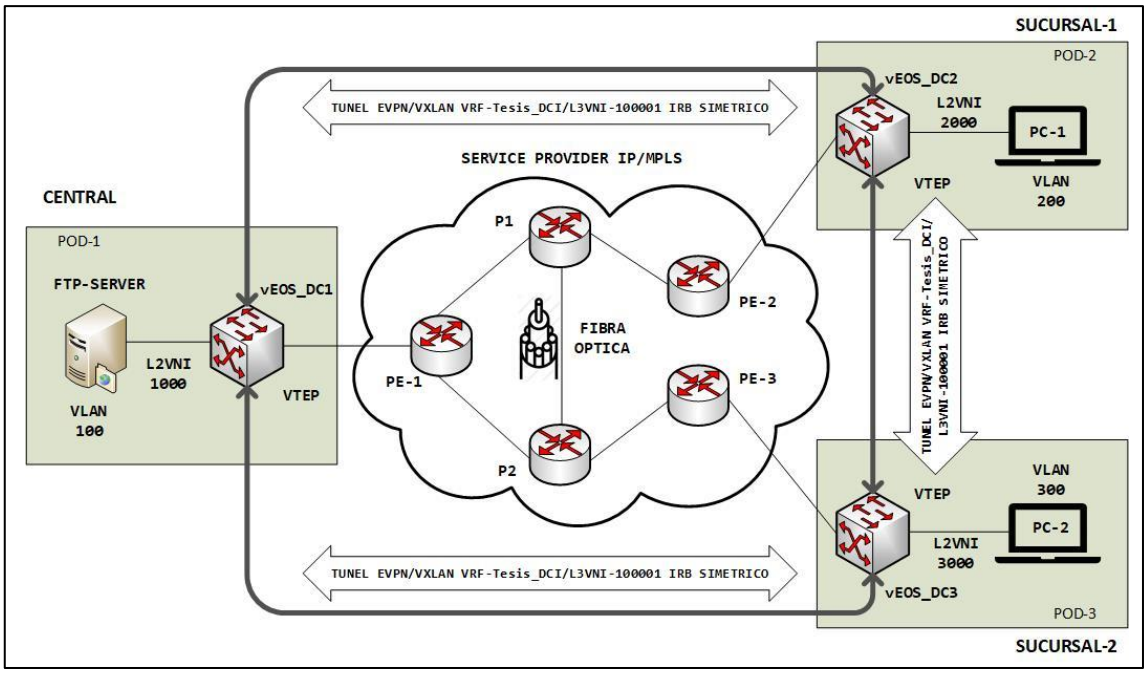

Figura 16-2: Formación de túneles EVPN/VXLAN a través de una instancia VRF-Tesis_DCI L3VNI 100001.

Realizado por: Mena Duval, 2021

En la figura 17-2, se muestra las rutas de tipo 5 (RT-5) de EVPN, el cual se conforma del prefijo IP y siguiente salto, aquí no se anuncia ninguna dirección MAC; ya que la característica principal de la ruta de tipo 5, es desacoplar los prefijos IP de las direcciones MAC.

```
CENTRAL(config)#show bgp evpn route-type ip-prefix ipv4
BGP routing table information for VRF default
Router identifier 10.0.0.1, local AS number 65001
Route status codes: s - suppressed, * - valid, > - active, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup
                    % - Pending BGP convergence
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

          Network                Next Hop              Metric  LocPref Weight  Path
 * >     RD: 10.0.0.1:1 ip-prefix 192.168.1.0/24
                                  -                     -       -       0       i
 * >     RD: 10.0.0.2:1 ip-prefix 192.168.2.0/24
                                  10.10.10.2            -       100     0       65002 i
 *       RD: 10.0.0.2:1 ip-prefix 192.168.2.0/24
                                  10.10.10.2            -       100     0       65003 65002 i
 * >     RD: 10.0.0.3:1 ip-prefix 192.168.3.0/24
                                  10.10.10.3            -       100     0       65003 i
 *       RD: 10.0.0.3:1 ip-prefix 192.168.3.0/24
                                  10.10.10.3            -       100     0       65002 65003 i
CENTRAL(config)#
```

Figura 17-2: Rutas de tipo 5, que contienen los prefijos IP y siguiente salto confeccionados por el plano de control MP-BGP EVPN.

Realizado por: Mena Duval, 2021

2.3.2.7. Pruebas de conectividad y análisis general

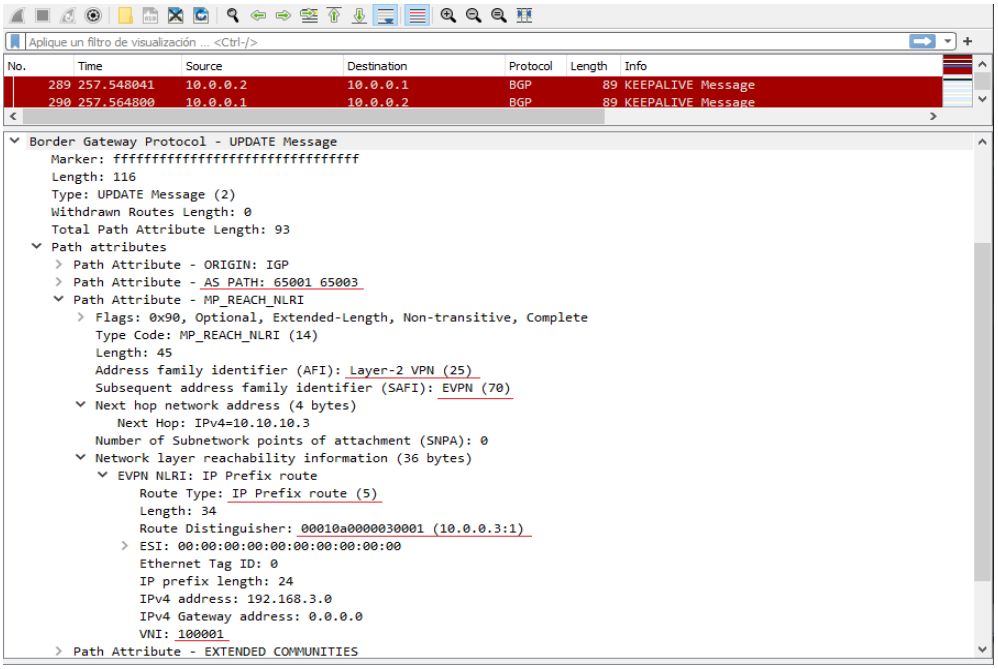

Figura 18-2: Establecimiento del peering BGP entre los ASN 65001-65003

Realizado por: Mena Duval, 2021

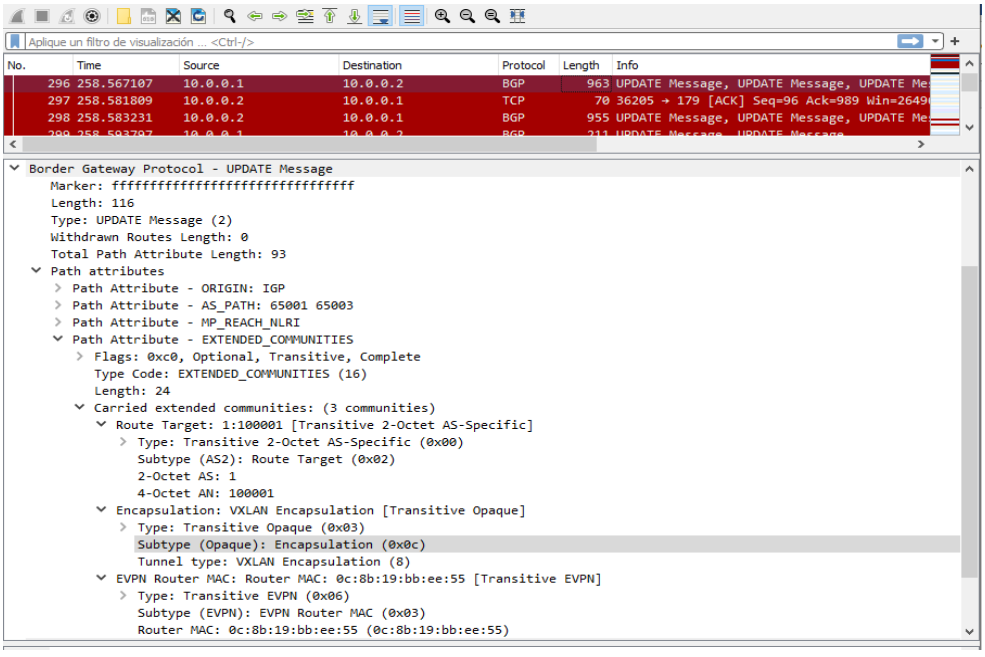

Figura 19-2: Parámetros de comunidades extendidas en el peering BGP

Realizado por: Mena Duval, 2021

En esta sección se pretende demostrar mediante pruebas de ping (ICMP), la conectividad de la red, adicional a esto exponer capturas de tráfico por medio de wireshark con el objetivo de mostrar ciertos parámetros que se consideran necesarios al establecer las relaciones de vecindad entre los

dispositivos que conforman la topología DCI. Como primer punto la figura 18-2, muestra el establecimiento de las sesiones eBGP (Peering BGP), entre el ASN 65001 y 65003, se rigen dentro de la familia de direcciones AFI/SAFI *l2vpn evpn*, esta a su vez trabaja por medio del tipo de ruta 5 mediante *L3VNI = 100001*; además se aprecia la forma como el *Route-Distinguisher (RD)*, antepone una serie de caracteres permitiendo que una dirección sea única a nivel mundial.

En la figura 20-2, se muestra la configuración de la interfaz eth0 de la PC-1, puesto que al utilizar tramas gigantes es necesario ajustar la MTU tanto en el servidor como en los clientes, luego se asigna la dirección IP y se procede a realizar la prueba de conectividad.

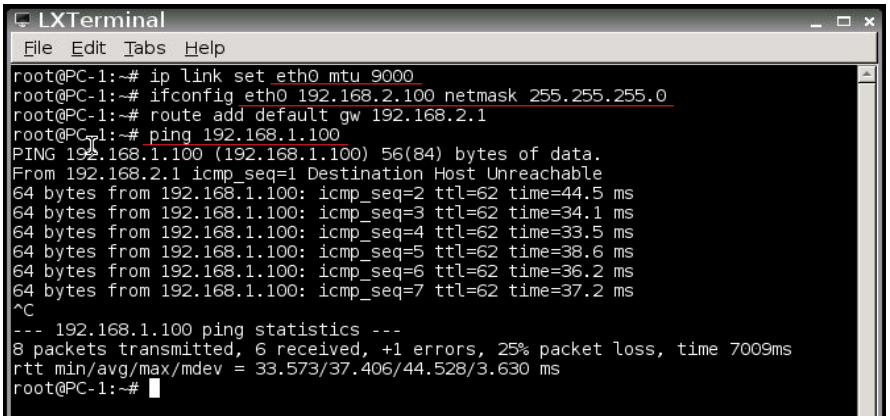

Figura 20-2: Prueba de conectividad mediante ping entre la PC-1 y el servidor FTP

Realizado por: Mena Duval, 2021

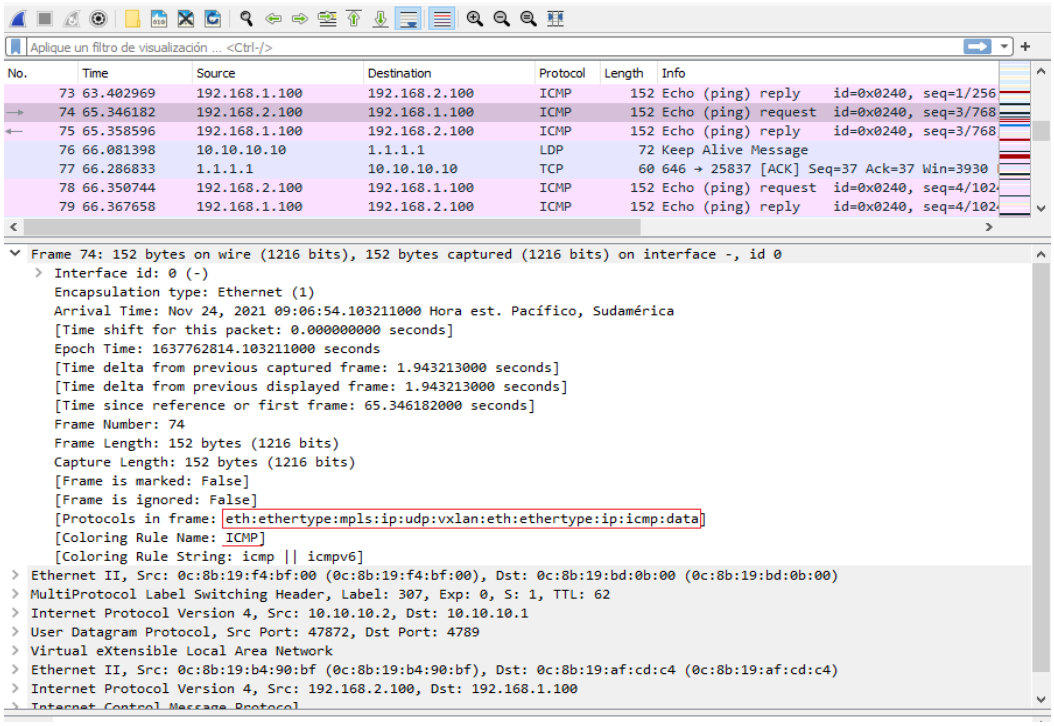

Figura 21-2: Formato de la trama ethernet completa

Realizado por: Mena Duval, 2021

La figura 21-2, muestra una prueba de ping entre la PC-1 y el servidor FTP; dentro de la captura se puede apreciar la variedad de protocolos que intervienen en la trama. Siguiendo la captura del mismo mensaje mostrado anteriormente, en la figura 22-2, dentro del protocolo IPv4 se puede apreciar que el bit DF (Don't fragment) esta seteado a 1, lo que significa que no se permite la fragmentación de paquetes; en si este es un requisito dentro de VXLAN, además se muestra el protocolo donde se realiza la encapsulación en este caso IP/UDP y las direcciones tanto del VTEP DC2 desde donde se envía el paquete ICMP (origen) y el VTEP DC1 que en este caso viene a ser el destino.

Dentro del protocolo UDP, el puerto de origen es seleccionado por cada uno de los VTEP, y el puerto de destino es el número 4789 que está establecido por la IANA, sin embargo, para otros propósitos aplicativos varias plataformas de enrutamiento permiten cambiar el número de puerto de destino. Para el transporte de los paquetes que viajaran a los diferentes centros de datos se establece el *L3VNI* 100001.

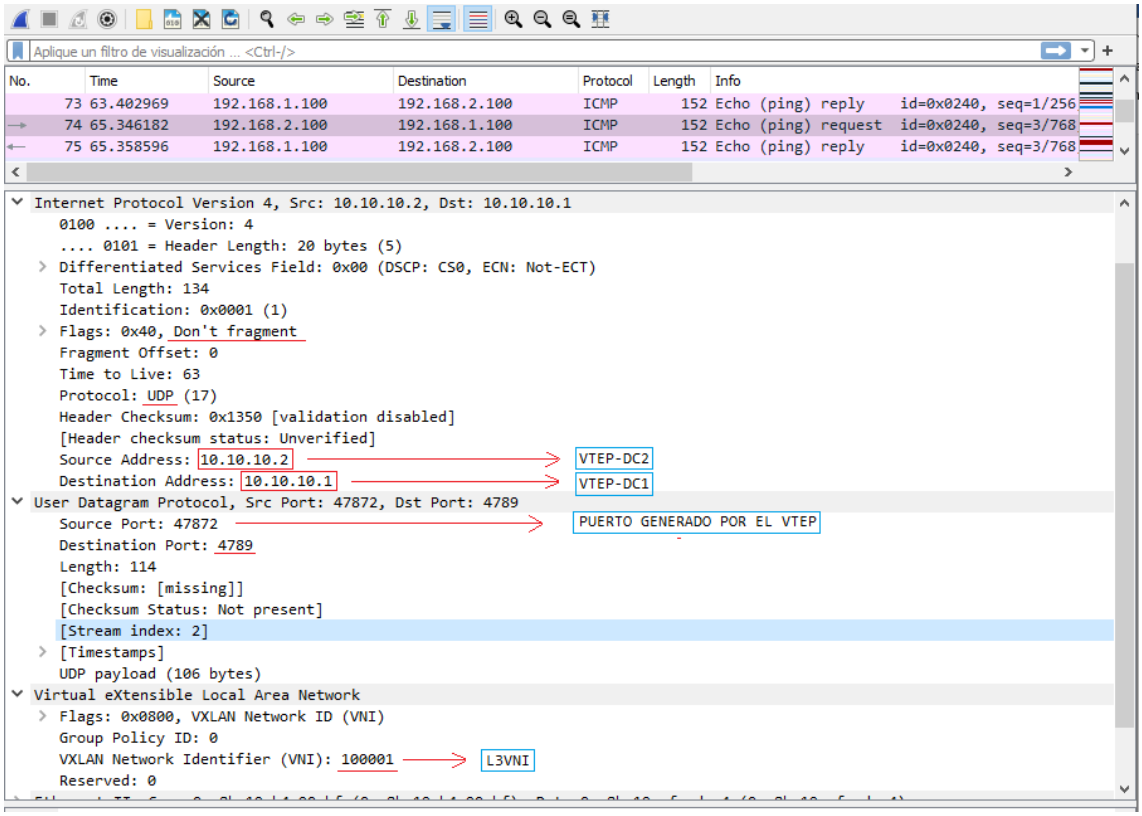

Figura 22-2: Protocolos que intervienen en la trama

Realizado por: Mena Duval, 2021

Ahora para demostrar que el tipo de enrutamiento y puenteo integrado (IRB), que se utiliza en la red es el *IRB simétrico*, simplemente se toma la captura de ping tanto para *echo request* y *echo reply*, con esto se puede constatar que el *VNI* que se utiliza tanto para la transmisión como para la recepción es el mismo, es decir, el *L3VNI = 100001*, esto se puede visualizar en la figura 23-2.

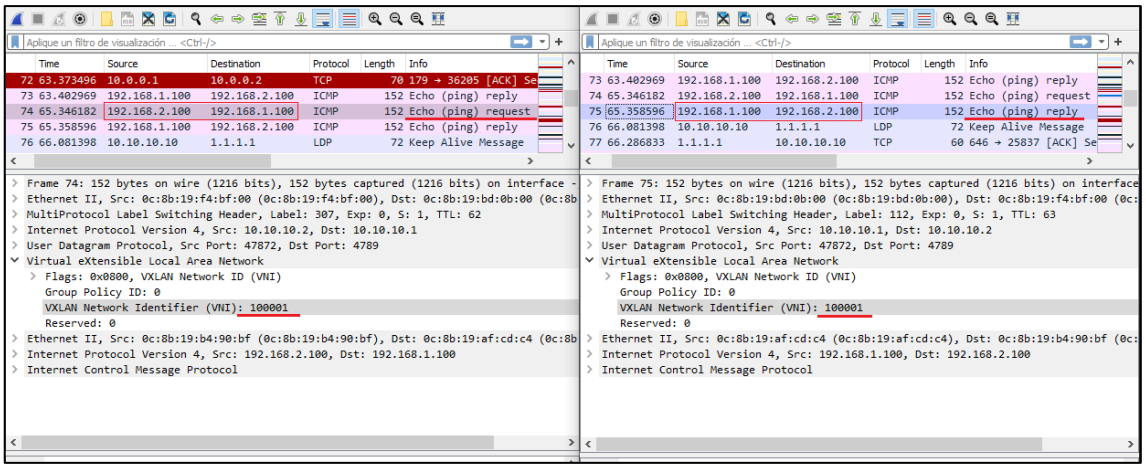

Figura 23-2: Uso de IRB simétrico dentro de la red DCI.

Realizado por: Mena Duval, 2021

2.4. Funcionamiento de la red comparativa con VLANs

La red que se describe a continuación dará paso a comparativas entre la red propuesta configurada con EVPN/VXLAN. La figura 24-2, muestra la configuración general de los conmutadores Arista que vienen representando a cada uno de los centros de datos que se mostraron previamente en la figura 2-2.

```
SUC-1(config)#ip routing
SUC-1(config)#int eth1
SUC-1(config-if-Et1)#no switchport
SUC-1(config-if-Et1)#ip add 10.10.0.6/30
SUC-1(config-if-Et1)#no shut
SUC-1(config-if-Et1)#exit
SUC-1(config)#vlan 200
SUC-1(config-vlan-200)#name v200
SUC-1(config-vlan-200)#exit
SUC-1(config)#int vlan 200
SUC-1(config-if-Vl200)#ip address 10.10.10.10/24
SUC-1(config-if-Vl200)#exit
SUC-1(config)#int eth3
SUC-1(config-if-Et3)#switchport mode access
SUC-1(config-if-Et3)#switchport access vlan 200
SUC-1(config-if-Et3)#no shut
SUC-1(config-if-Et3)#exit
SUC-1(config)#router ospf 10
SUC-1(config-router-ospf)#router-id 10.0.0.2
SUC-1(config-router-ospf)#log-adjacency-changes
SUC-1(config-router-ospf)#network 10.10.0.4/30 area 0
SUC-1(config-router-ospf)#network 10.10.10.0/24 area 0
SUC-1(config-router-ospf)#exit
SUC-1(config)#
```

Figura 24-2: Configuración de enrutamiento OSPF con su respectiva VLAN

Realizado por: Mena Duval, 2021

Como se puede observar en la figura 24-2, la configuración de cada uno de los centros de datos se apoya del enrutamiento IP tradicional con OSPF para su salida hacia el mundo exterior, para

este caso no es necesario la aplicación de tramas gigantes por tanto no se modifica el mtu predeterminado de 1500 bytes, por último la configuración de cada uno de los dispositivos en el proveedor de servicios es la misma que se presentó en el apartado 2.3.1., con un pequeño ajuste en el MTU que fue asignado a 1504 bytes debido a la cabecera de MPLS LDP y conserva el esquema de direccionamiento expuesto en la tabla 4-2, sin embargo, dentro de cada centro de datos el único punto que difiere con el direccionamiento establecido son las redes en las que se encuentran las diferentes VLANs, es así que el nuevo direccionamiento para las VLANs se muestra en la tabla 4-3.

Tabla 4-3: Esquema de direccionamiento de VLANs para la red comparativa.

Equipo	*Interfaz*	*Dirección IP*	*Máscara de red*	*Gateway*	*VLAN*
Arista DC1-POD1 Central	Interfaz VLAN	192.168.1.10	255.255.255.0		100
Arista DC1-POD1 Central	Interfaz VLAN	10.10.10.10	255.255.255.0		200
Arista DC1-POD1 Central	Interfaz VLAN	172.16.1.10	255.255.255.0		300
FTP-SERVER	eth0	192.168.1.100	255.255.255.0	192.168.1.10	100
PC-1	eth0	10.10.10.100	255.255.255.0	10.10.10.10	200
PC-2	eth0	172.16.1.100	255.255.255.0	172.16.1.10	300

Realizado por: Mena Duval, 2021

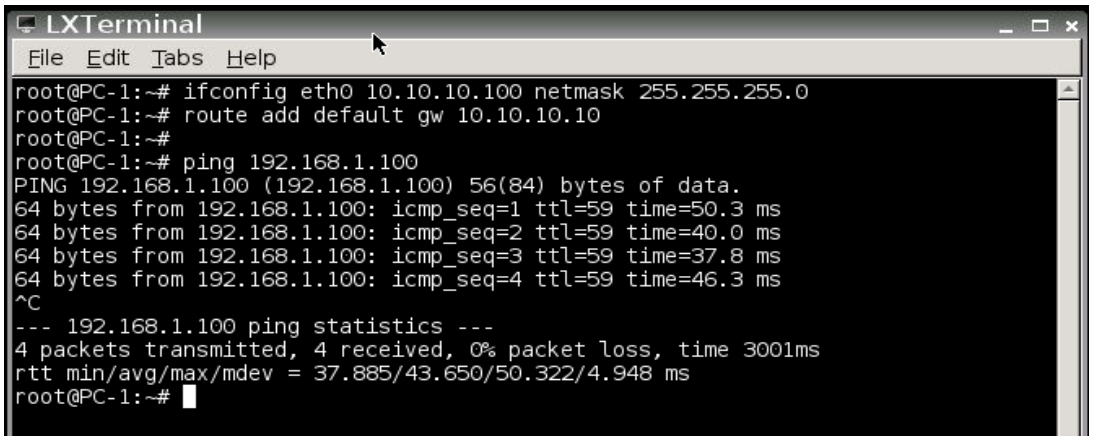

Figura 25-2: Prueba de conectividad

Realizado por: Mena Duval, 2021

La figura 25-2, muestra una prueba de conectividad entre el centro de datos denominado Sucursal-1 que aloja a la Vlan 200 hacia el centro de datos Central que aloja a la Vlan 100.

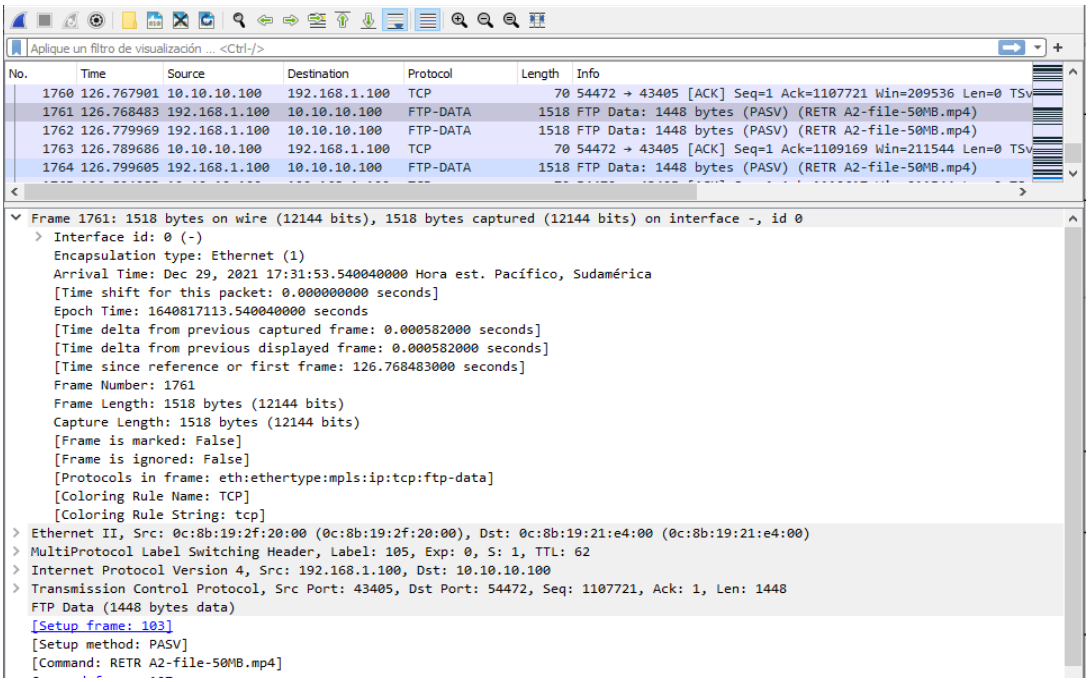

Figura 26-2: Captura de tráfico para la red comparativa mediante wireshark

Realizado por: Mena Duval, 2021

La figura 26-2, muestra una captura por medio de wireshark en la cual se puede apreciar todos los protocolos que conforman la estructura de la trama de datos, de esta forma los paquetes que salen por cada centro de datos viajan con una mtu de 1500 bytes hacia la red de transporte MPLS hasta llegar a su respectivo destino.

2.5. Método de generación de tráfico

La forma como se va a generar el tráfico dentro de la red, es por medio de descargas desde un servidor con el sistema operativo Ubuntu 20.04 LTS (Focal Fossa), configurado con el protocolo *Vsftpd (Very Secure FTP Daemon)*, que se considera como uno de los servidores FTP más completos y potentes de la distribución Linux, se hace uso de FTP ya que es el método con mayor similitud en referencia a la forma en cómo operan los centros de datos (Fosco Connect, 2012); el servidor va a estar alojado en el centro de datos denominado *CENTRAL* y los centros de datos *Sucursal-1 y Sucursal-2*, alojan a los clientes, estos a su vez contienen el programa *Filezilla* por medio del cual realizarán la conexión con el servidor *vsftpd* y procederán a realizar las descargas. La figura 27-2, muestra la configuración del servidor *Vsftpd.*

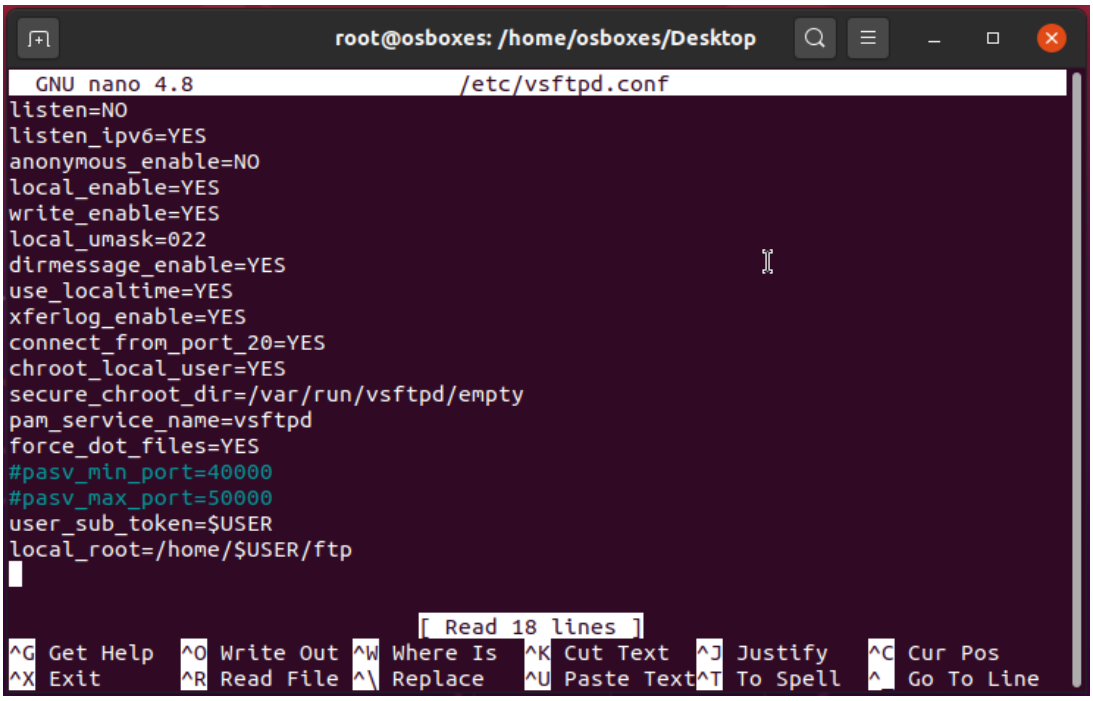

Figura 27-2: Configuración del servidor Vsftpd en Ubuntu 20.04

Realizado por: Mena Duval, 2021

Se han creado dos usuarios que permitirán las descargar de archivos, uno por cada sucursal, los usuarios se pueden apreciar en la figura 28-2.

```
ftp:x:126:133:ftp daemon,,,:/srv/ftp:/usr/sbin/nologin
sshd:x:127:65534::/run/sshd:/usr/sbin/nologin
user1:x:1001:1001:Andres,,,:/home/user1:/bin/bash
user2:x:1002:1002:Sara,,,:/home/user2:/bin/bash
(END)
```

Figura 28-2: Usuarios del servidor Vsftpd

Realizado por: Mena Duval, 2021

Dentro de cada usuario se alojan tres archivos que tienen los siguientes tamaños: 1 MB (carga baja), 50 MB (carga media) y 100 MB (carga alta), estos se muestran en la figura 28-2 (Kiran, 1998, pp.39-40) (Adusei, 2016, pp.64-65).

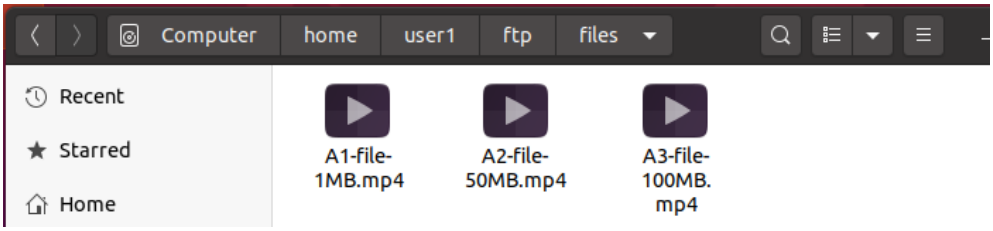

Figura 29-2: Archivos contenidos dentro de cada usuario para la descarga por medio de FTP

Realizado por: Mena Duval, 2021

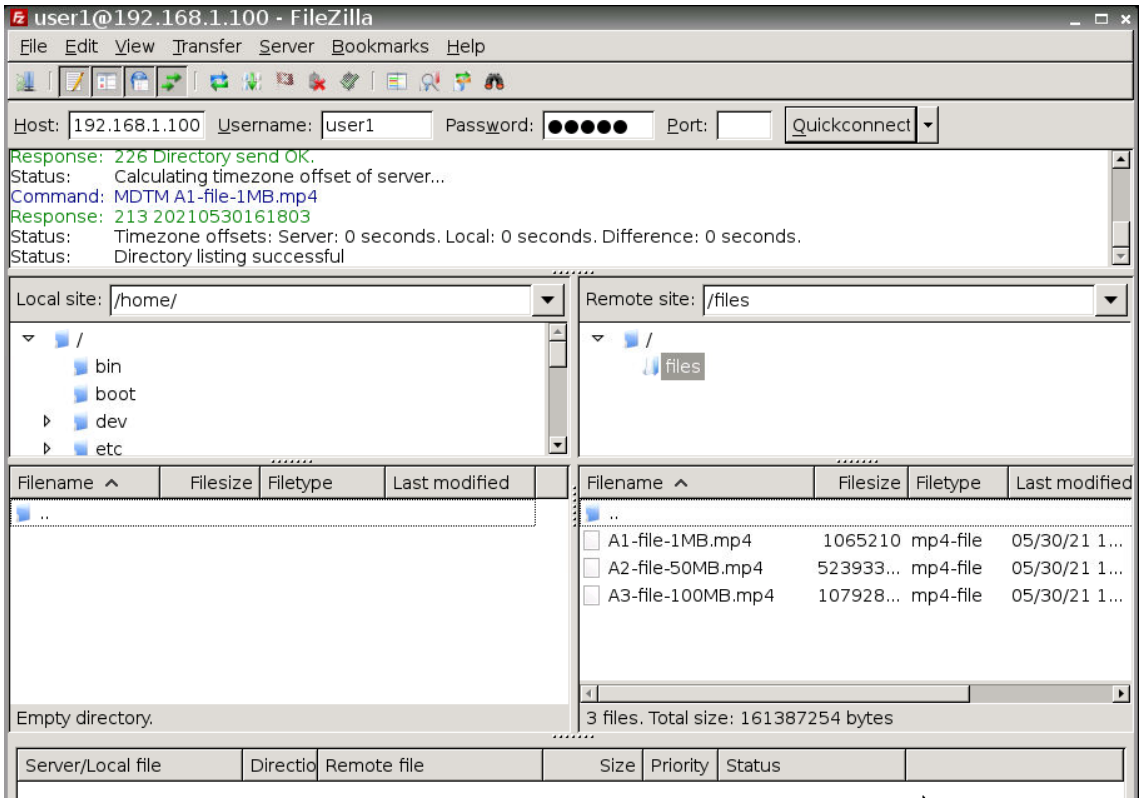

Figura 30-2: Vista de los archivos desde uno de los clientes por medio de FileZilla.

Realizado por: Mena Duval, 2021

2.6. Obtención de parámetros de rendimiento

Dentro de la topología propuesta para la recolección de los parámetros de tráfico, se emplea la herramienta Wireshark, esta a su vez trabaja de forma predeterminada junto con Gns3. Las capturas se realizan dentro del proveedor de servicios y en las conexiones directas a los centros de datos (conmutadores Arista - PE), el tiempo de recolección de los datos, varía de acuerdo al tamaño de los paquetes y sumado a esto el ancho de banda que permite cada uno de los dispositivos que forman parte de la red. El ancho de banda de los dispositivos están limitados por los propios creadores por motivos de licencias, de esta forma los valores de ancho de banda promedio que se puede alcanzar tanto con los conmutadores Arista como con los enrutadores cisco es de $\pm\ 1 \sim 2\ [Mbps]$ sin licenciamiento (Nouri, 2017; Gns3, 2019). La figura 31-2, muestra una captura del tráfico en base a una descarga desde el cliente que se encuentra en la *Sucursal-1* hacia el servidor FTP.

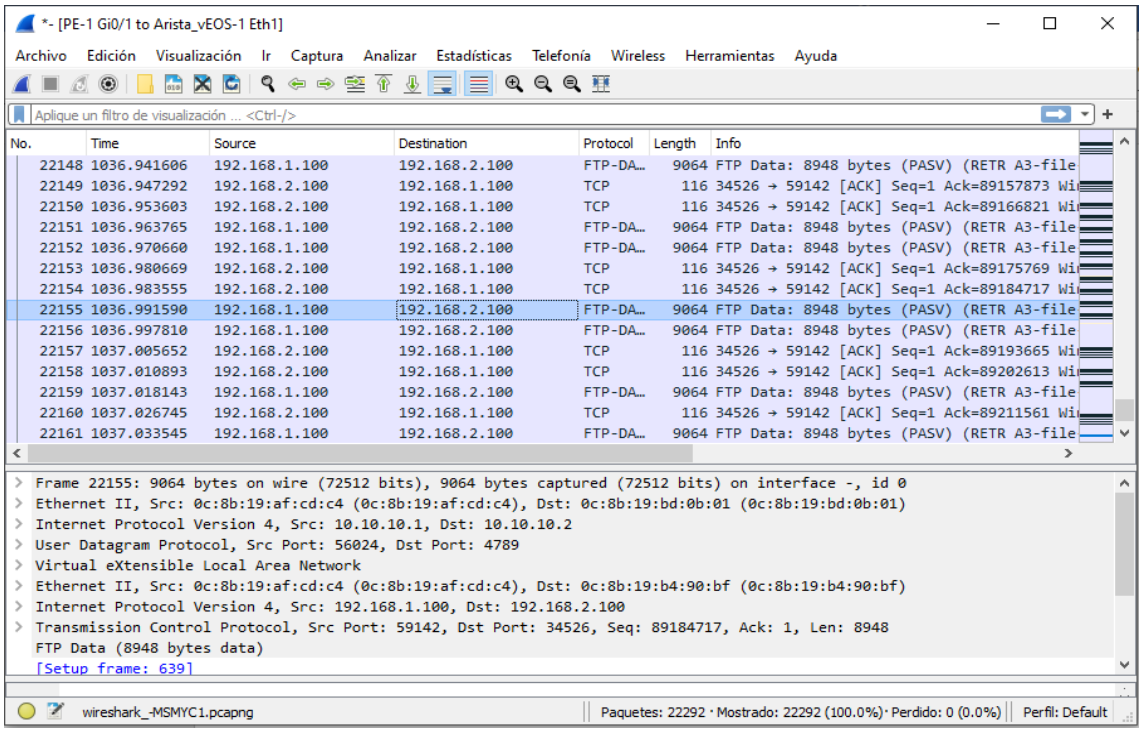

Figura 31-2: Captura de tráfico por medio de Wireshark

Realizado por: Mena Duval, 2021

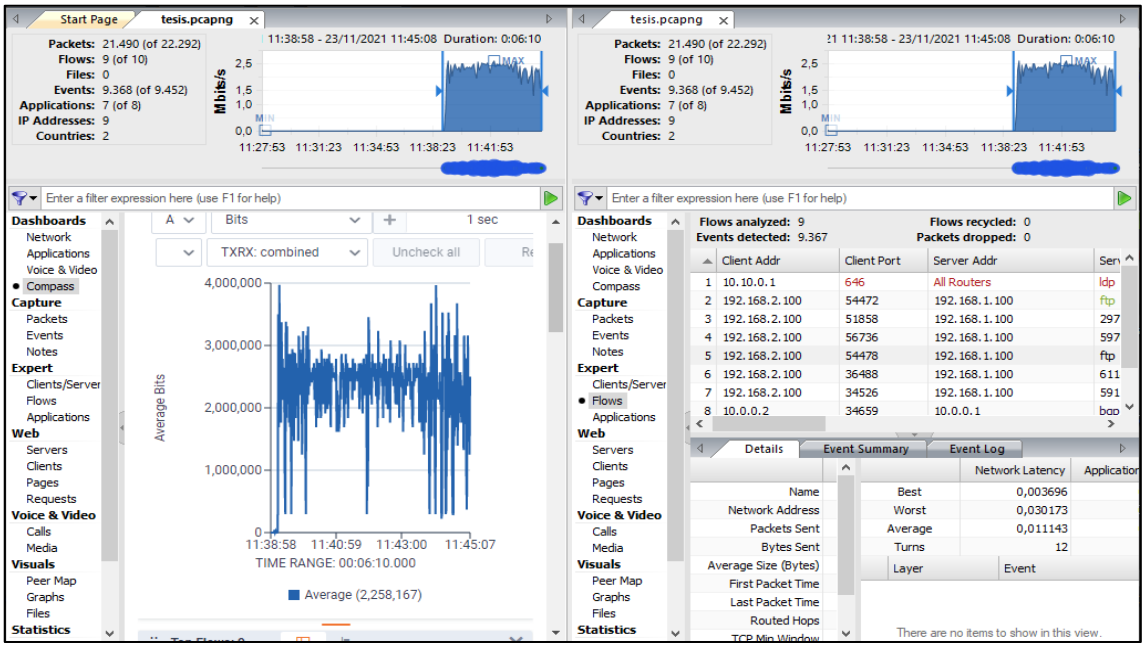

Figura 32-2: Interfaz de Omnipeek, extracción de ancho de banda y latencia.

Realizado por: Mena Duval, 2021

Para la obtención de los parámetros de rendimiento (ancho de banda, latencia y perdida de paquetes), se hace uso de un analizador de protocolos de red denominado *Omnipeek*, gracias a los archivos (*.pcap)* generados por wireshark se pueden importar dentro de la aplicación omnipeek

para su posterior análisis. La figura 32-2, muestra una de las capturas de tráfico procesadas por omnipeek y la forma como se extrae el ancho de banda y latencia.

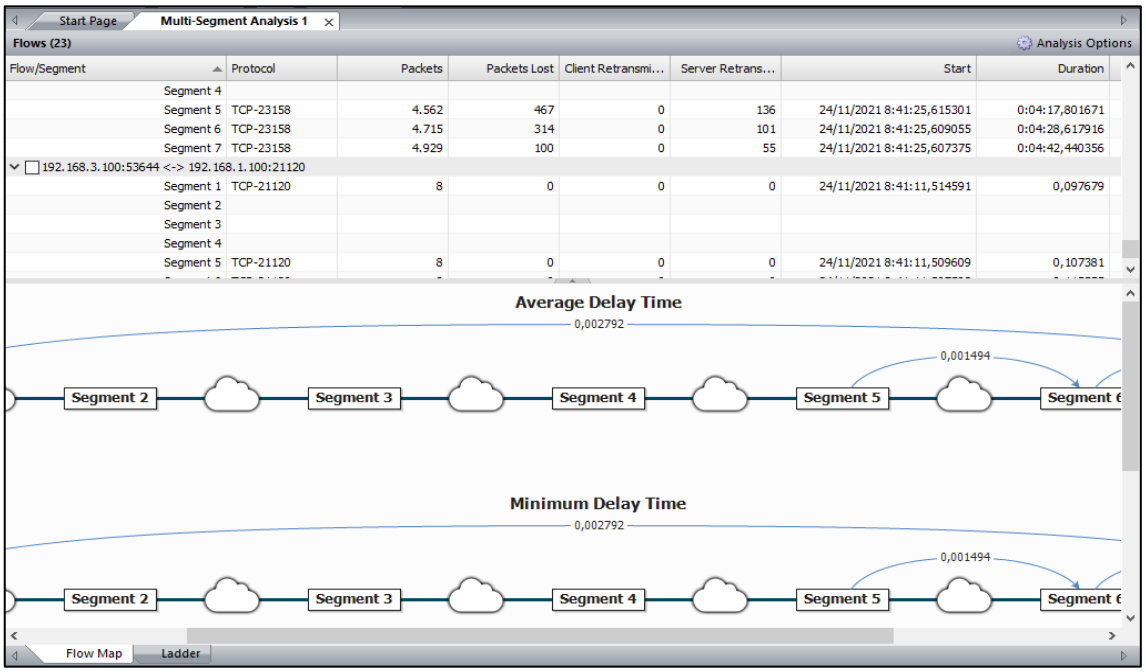

Figura 33-2: Cálculo de pérdida de paquetes en omnipeek

Realizado por: Mena Duval, 2021

Para la obtención de la perdida de paquetes, omnipeek posee una función denominada *Multi-segment Analysis*, donde permite analizar por segmentos para estimar la perdida de paquetes que se generó dentro de la red, esto se muestra en la figura 33-2.

Para los valores de *Jitter* fue necesario realizar un análisis externo, sin embargo, omnipeek brinda un valor base para poder calcular el jitter este se denomina *Delta time* que hace referencia al tiempo de ida y vuelta (RTT), es decir, el *delta time* refleja el tiempo en [*ms*] entre los paquetes de datos que se envían y reciben a través de una red; en términos generales el jitter es la variación del *Round-trip time*. Es por esto que se realizó los cálculos de forma manual a través del programa de cálculo *Excel*, donde se tomaron los valores de RTT y se procedió a calcular la varianza y desviación estándar para así poder obtener las variaciones de tiempo (*jitter*) en [*ms*] de los paquetes que se enviaron y recibieron entre los clientes y el servidor. La ecuación 1-2, muestra la forma como se calculó el jitter.

$$Jitter\ FTP = \sqrt{\frac{\sum_1^n (RTT_i - \overline{RTT})^2}{n}} * 1000\ [ms]$$

Ecuación 1-2: Cálculo del jitter de FTP

CAPÍTULO III

3. RESULTADOS Y DISCUSIÓN DE RESULTADOS

En este capítulo se aborda el análisis de resultados en función de los parámetros de rendimiento obtenidos en base a descargas realizadas a través de los centros de datos denominados *Sucursal-1 (PC-1)* y *Sucursal-2 (PC-2)*, hacia el centro de datos denominado *Central* que contiene el servidor FTP, como se describió en el capítulo anterior sección 2.3.3., el servidor FTP tiene creado 2 usuarios y estos a su vez poseen en su almacenamiento tres archivos (*1[MB]-low load, 50[MB]-medium load, 100[MB]-High load*). Se procedió a realizar un número total de 50 descargas por cada paquete que contenían los usuarios, para poder establecer los valores promedio de los parámetros de rendimiento de la red al interconectar centros de datos mediante los protocolos *BGP EVPN/VXLAN* y mediante enrutamiento IP junto a VLANs.

3.1. Parámetros de rendimiento recomendados por estándares internacionales

Tomando como referencia los estándares UIT-T (Rec. G.1010, Rec. Y.1541) e IEEE 802.1p, Buñay (2013, pp.74-83), menciona y establece valores recomendados en función de dichos estándares a los cuales los parámetros de rendimiento (latencia, jitter, perdida de paquetes) deben apegarse con la finalidad de garantizar que la red sea eficiente y se asegure la calidad del servicio a los puntos finales.

3.1.1. Valores recomendados de latencia

Tabla 1-3: Valores de latencia recomendados por los estándares internacionales

Clase de calidad de servicio	Descripción	Umbral de retardo internacional (ms)
0	Tiempo real, alta interacción, sensibles al retardo (voz y video en tiempo real)	100
1	Tiempo real, interactivos, sensibles al retardo (voz y video en tiempo real de menor calidad)	150
2	Datos de alta prioridad (transaccionales, altamente interactivos)	200
3	Datos de mediana prioridad (Datos transaccionales interactivos)	225
4	Datos de baja prioridad (transacciones cortas, datos en grandes cantidades, flujo continuo de video streaming)	250
5	Datos de mejor esfuerzo	300

Fuente: (Buñay, 2013, p.75)

Realizado por: Mena Duval, 2021

La tabla 1-3, muestra diversas clases de servicios con sus respectivos valores de latencia recomendados, para efectos de este trabajo se toma la clase de servicio número 4, en función de datos en grandes cantidades que es el enfoque que manejan las descargas desde el servidor FTP. En base a la tabla anterior, se establece las siguientes categorías para los valores de latencia:

Tabla 2-3: Categorización de la latencia

Categoría de Latencia	*Latencia*
Excelente	< 150 ms
Adecuado	< 250 ms
Conforme	< 350 ms
No Adecuado	< 450 ms

Fuente: (Hafiz y Susianto, 2019, p.3)

Realizado por: Mena Duval, 2021

3.1.2. Valores permisibles de perdida de paquetes

Tabla 3-3: Valores permisibles para pérdida de paquetes recomendados por los estándares internacionales

Clase de calidad de servicio	Descripción	Umbral de pérdida de paquetes (%)
0	Tiempo real, alta interacción, sensibles al retardo (voz y video en tiempo real)	1%
1	Tiempo real, interactivos, sensibles al retardo (voz y video en tiempo real de menor calidad)	3%
2	Datos de alta prioridad (transaccionales, altamente interactivos)	3%
3	Datos de mediana prioridad (Datos transaccionales interactivos)	5%
4	Datos de baja prioridad (transacciones cortas, datos en grandes cantidades, flujo continuo de video streaming)	5%
5	Datos de mejor esfuerzo	5%

Fuente: (Buñay, 2013, p.82)

Realizado por: Mena Duval, 2021

Tabla 4-3: Categorización de los valores porcentuales permisibles para pérdida de paquetes

Degradación de la perdida de paquetes	***% Perdida de Paquetes***
Excelente	< 1 %
Adecuado	> 1 y < 3 %
Conforme	>3 y <5 %
No Adecuado	> 5 %

Fuente: (Buñay, 2013, p.82)

Realizado por: Mena Duval, 2021

3.1.3. Valores recomendados de jitter

Tabla 5-3: Valores de jitter recomendados por los estándares internacionales

Clase de calidad de servicio	Descripción	Umbral de jitter internacional (ms)
0	Tiempo real, alta interacción, sensibles al retardo (voz y video en tiempo real)	45
1	Tiempo real, interactivos, sensibles al retardo (voz y video en tiempo real de menor calidad)	50
2	Datos de alta prioridad (transaccionales, altamente interactivos)	55
3	Datos de mediana prioridad (Datos transaccionales interactivos)	N/A
4	Datos de baja prioridad (transacciones cortas, datos en grandes cantidades, flujo continuo de video streaming)	N/A
5	Datos de mejor esfuerzo	N/A

Fuente: (Buñay, 2013, p.78)

Realizado por: Mena Duval, 2021

Para los datos en grandes cantidades, los estándares no presentan como tal una valoración para el jitter, puesto que las aplicaciones que son enfocadas al uso masivo de datos no se ven afectados de forma critica con los valores de jitter en comparación con aplicaciones que se manejan en tiempo real como por ejemplo las llamadas VoIP; Lewis et al. (2006, p.171), califican como less-than-best-effort (menos que mejor esfuerzo) a las aplicaciones basadas en FTP, backups, y aplicaciones no críticas. A continuación, en la tabla 6-3, en base a los estándares ITU-T Rec. G.1011 e ITU-T Rec. E.800 se presenta una clasificación en función de la relevancia que tienen los parámetros de rendimiento dentro de las diferentes aplicaciones; donde "-" hace referencia a menos relevante y "+" a más relevante.

Tabla 6-3: Relevancia de los parámetros de rendimiento dentro de las diversas aplicaciones de datos

Aplicación	Velocidad de transmisión		Latencia	Jitter	Pérdida de Paquetes
	Downstream	Upstream			
Texto	++	-	++	-	+++
Multimedia	+++	-	++	+	+++
Descarga de Archivos	+++	-	+	-	+
Transacciones	-	-	++	-	+++
Streaming	+++	-	+	-	+
Voip	+	+	+++	+++	+
Gamming	+	+	+++	++	+++

Fuente: (ITU, 2017, pp.59-60)

Realizado por: Mena Duval, 2021

Como se puede observar en la tabla 6-3, dentro de la descarga de archivos los parámetros con relevancia son la latencia, ancho de banda (downstream) y la perdida de paquetes.

3.2. Resultados con BGP EVPN/VXLAN

3.2.1. Parámetros de rendimiento en base a paquetes de 1 MB

La tabla 7-3, muestra los resultados de parámetros de rendimiento promedios, obtenidos en base a 50 descargas de paquetes de 1 MB, que se transportaban por medio de los diferentes enlaces que conectan a los centros de datos a través de un proveedor de servicios configurado con MPLS.

Tabla 7-3: Resultados de rendimiento para descargas de paquetes de 1 MB

Enlace DC1-PE1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
1057 Kb/s	14.63 ms	187.71 ms	2.71 %
Enlace PE1-P1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**

689.77 Kb/s	22.37 ms	161.24 ms	2.91 %
Enlace P1-PE2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
675 Kb/s	31 ms	155.56 ms	2.94 %
Enlace PE2-DC2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
673,68 Kb/s	39.60 ms	170.60 ms	2.91 %
Enlace PE1-P2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
691,89 Kb/s	23.67 ms	160.09 ms	2.83 %
Enlace P2-PE3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
687,46 Kb/s	32.67 ms	169.10 ms	2.88 %
Enlace PE3-DC3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
683,26 Kb/s	41.83 ms	165.50 ms	2.96 %

Realizado por: Mena Duval, 2021

3.2.2. Parámetros de rendimiento en base a paquetes de 50 MB

La tabla 8-3, muestra los resultados de parámetros de rendimiento promedios, obtenidos en base a 50 descargas de paquetes de 50 MB, que se transportaban por medio de los diferentes enlaces que conectan a los centros de datos a través de un proveedor de servicios configurado con MPLS.

Tabla 8-3: Resultados de rendimiento para descargas de paquetes de 50 MB

Enlace DC1-PE1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
1235.14 Kb/s	14.34 ms	94 ms	2.06 %
Enlace PE1-P1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
692.34 Kb/s	23.36 ms	177.80 ms	2.78 %

Enlace P1-PE2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
678.03 Kb/s	32.82 ms	180.16 ms	2.79 %
Enlace PE2-DC2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
681,89 Kb/s	41.95 ms	191.03 ms	2.85 %
Enlace PE1-P2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
685,42 Kb/s	23.81 ms	179.87 ms	2.77 %
Enlace P2-PE3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
670,87 Kb/s	32.60 ms	182.08 ms	2.79 %
Enlace PE3-DC3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
669,78 Kb/s	41.23 ms	191.93 ms	2.84 %

Realizado por: Mena Duval, 2021

3.2.3. Parámetros de rendimiento en base a paquetes de 100 MB

La tabla 9-3, muestra los resultados de parámetros de rendimiento promedios, obtenidos en base a 50 descargas de paquetes de 100 MB, que se transportaban por medio de los diferentes enlaces que conectan a los centros de datos a través de un proveedor de servicios configurado con MPLS.

Tabla 9-3: Resultados de rendimiento para descargas de paquetes de 100 MB

Enlace DC1-PE1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
1277.41 Kb/s	15.57 ms	83.39 ms	2.03 %
Enlace PE1-P1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
694.99 Kb/s	24.42 ms	175.54 ms	2.56 %
Enlace P1-PE2			

Ancho de banda	Latencia	Jitter	Pérdida de paquetes
682.59 Kb/s	33.50 ms	176.92 ms	2.58 %
Enlace PE2-DC2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
674,33 Kb/s	41.86 ms	189.69 ms	2.63 %
Enlace PE1-P2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
687,66 Kb/s	25.62 ms	186.33 ms	2.54 %
Enlace P2-PE3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
676,57 Kb/s	34.78 ms	187.47 ms	2.58 %
Enlace PE3-DC3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
673,60 Kb/s	43.46 ms	199.65 ms	2.64 %

Realizado por: Mena Duval, 2021

3.3. Resultados con VLANs y enrutamiento IP

3.3.1. Parámetros de rendimiento en base a paquetes de 1 MB

La tabla 10-3, muestra los resultados de parámetros de rendimiento promedios, obtenidos en base a 50 descargas de paquetes de 1 MB, que se transportaban por medio de los diferentes enlaces que conectan a los centros de datos a través de un proveedor de servicios configurado con MPLS.

Tabla 10-3: Resultados de rendimiento para descargas de paquetes de 1 MB

Enlace DC1-PE1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
170.55 Kb/s	12.41 ms	206.82 ms	3.00 %
Enlace PE1-P1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
89.50 Kb/s	22.61 ms	249.52 ms	3.24 %

Enlace P1-PE2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
88.88 Kb/s	30.67 ms	260.35 ms	3.28 %
Enlace PE2-DC2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
91,30 Kb/s	39.37 ms	257.94 ms	3.27 %
Enlace PE1-P2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
88,95 Kb/s	22.57 ms	281.12 ms	3.19 %
Enlace P2-PE3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
88,79 Kb/s	30.96 ms	274.08 ms	3.23 %
Enlace PE3-DC3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
90.37 Kb/s	39.94 ms	286.81 ms	3.32 %

Realizado por: Mena Duval, 2021

3.3.2. Parámetros de rendimiento en base a paquetes de 50 MB

La tabla 11-3, muestra los resultados de parámetros de rendimiento promedios, obtenidos en base a 50 descargas de paquetes de 50 MB, que se transportaban por medio de los diferentes enlaces que conectan a los centros de datos a través de un proveedor de servicios configurado con MPLS.

Tabla 11-3: Resultados de rendimiento para descargas de paquetes de 50 MB

Enlace DC1-PE1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
184.78 Kb/s	13.38 ms	103.23 ms	2.92 %
Enlace PE1-P1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
97.66 Kb/s	21.79 ms	233.34 ms	3.08 %
Enlace P1-PE2			

Ancho de banda	Latencia	Jitter	Pérdida de paquetes
96.58 Kb/s	30.02 ms	241.90 ms	3.16 %
Enlace PE2-DC2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
95.71 Kb/s	38.50 ms	260.03 ms	3.20 %
Enlace PE1-P2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
98.01 Kb/s	21.67 ms	227.11 ms	3.07 %
Enlace P2-PE3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
96,74 Kb/s	30.51 ms	228.67 ms	3.16 %
Enlace PE3-DC3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
95,93 Kb/s	39.11 ms	250.36 ms	3.20 %

Realizado por: Mena Duval, 2021

3.3.3. Parámetros de rendimiento en base a paquetes de 100 MB

La tabla 12-3, muestra los resultados de parámetros de rendimiento promedios, obtenidos en base a 50 descargas de paquetes de 100 MB, que se transportaban por medio de los diferentes enlaces que conectan a los centros de datos a través de un proveedor de servicios configurado con MPLS.

Tabla 12-3: Resultados de rendimiento para descargas de paquetes de 100 MB

Enlace DC1-PE1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
180.06 Kb/s	14.07 ms	104.79 ms	2.73 %
Enlace PE1-P1			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
96.31 Kb/s	22.66 ms	216.00 ms	2.80 %
Enlace P1-PE2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**

95.10 Kb/s	31.56 ms	223.73 ms	2.85 %
Enlace PE2-DC2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
94.30 Kb/s	40.36 ms	235.33 ms	2.88 %
Enlace PE1-P2			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
95.35 Kb/s	22.82 ms	221.80 ms	2.79 %
Enlace P2-PE3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
94.21 Kb/s	31.31 ms	225.33 ms	2.83 %
Enlace PE3-DC3			
Ancho de banda	**Latencia**	**Jitter**	**Pérdida de paquetes**
93.27 Kb/s	39.20 ms	239.57 ms	2.85 %

Realizado por: Mena Duval, 2021

3.4. Comparación de los resultados obtenidos

En base a los resultados presentados en las tablas anteriores se procedió a calcular los parámetros de rendimiento promedios generales de toda la red, lo que permitirá comparar y establecer que tecnología obtiene mejores resultados cuando la red está totalmente operativa y sus enlaces están congestionados de datos, prosiguiendo en los apartados siguientes se han representado gráficamente cada uno de estos parámetros para distinguir de mejor manera la diferencia que hay entre cada uno de ellos.

3.4.1. Comparación entre Ancho de banda

Como se puede observar en el grafico 1-3, el ancho de banda para EVPN-VxLAN supera los 700 Kbps y es notablemente mayor que el ancho de banda en el escenario con Vlans que apenas supera los 100 Kbps, esto es debido a que el uso de tramas gigantes como requisito para el funcionamiento de VxLAN en centros de datos, mejora el ancho de banda de la red; sin embargo, como se puede notar el ancho de banda en ambos escenarios no alcanza la velocidad de 1 Mbps que ofrece la plataforma de simulación de redes Gns3, esto debido a que cada uno de los enlaces

han sido configurados de forma tal que la red sea lo más parecido a una red real con un proveedor de servicios MPLS y el medio físico con fibra óptica (ver apartado 2.2.2.).

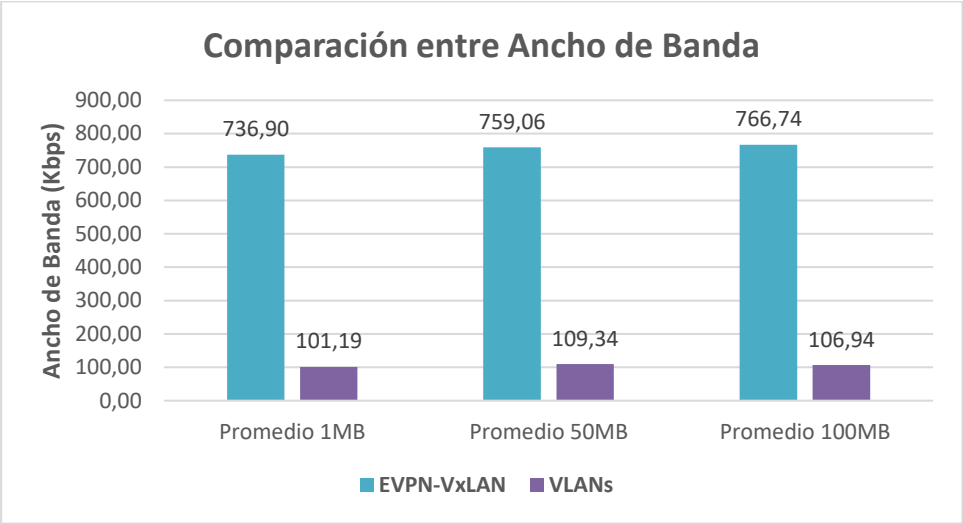

Gráfico 1-3: Comparación del Ancho de banda promedio de los escenarios propuestos

Realizado por: Mena Duval, 2021

Además, en base a la tabla 6-3, tomando en cuenta que uno de los parámetros esenciales dentro de las descargas de archivos es el downstream, de esta forma con el uso del jumbo frames para VxLAN se puede obtener un ancho de banda mayor y así lograr la optimización de la red. La tabla 13-3, muestra los tiempos aproximados de cada uno de los paquetes

3.4.2. Comparación entre Latencias

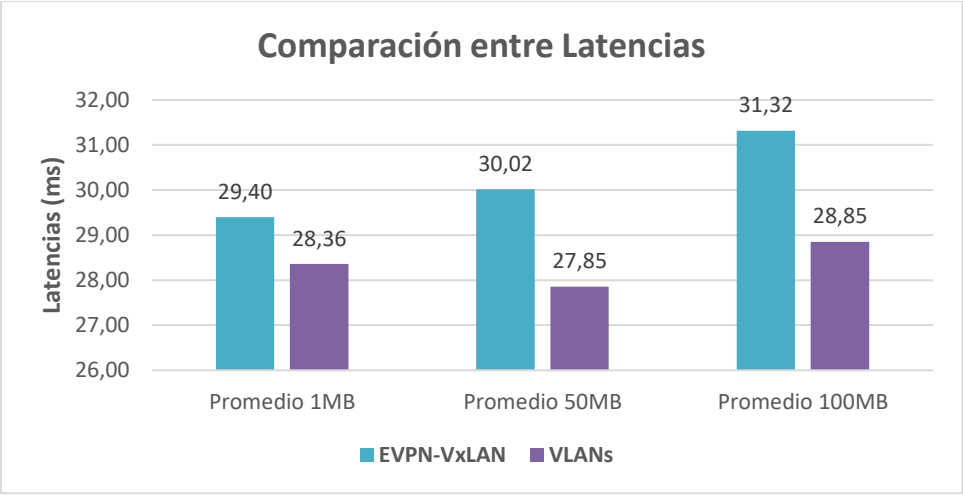

Gráfico 2-3: Comparación entre latencias promedio de los escenarios propuestos

Realizado por: Mena Duval, 2021

En el grafico 2-3, se muestran los valores de latencias promedio generales de los escenarios propuestos, en este caso la latencia con Vlans es menor en comparación a los valores de latencias generados en el escenario con EVPN/VxLAN, esto se debe a la sobrecarga que genera el encabezado VxLAN de 50 bytes, por tal motivo los equipos deben procesar un mayor número de paquetes incorporados en la trama y esto conlleva un uso mayor de recursos tanto de CPU como de memoria lo que provoca tal degradación de la latencia.

Tomando como referencia la tabla 2-3, se procede a realizar la categorización de los valores de latencias obtenidos en función de los dos escenarios propuestos como se muestra en la tabla 13-3.

Tabla 13-3: Categorización de latencias

Categorización de los parámetros de Latencia					
	EVPN-VxLAN (ms)			VLANs (ms)	
Carga baja 1MB	29,40	Excelente		28,36	Excelente
Carga media 50MB	30,02	Excelente		27,85	Excelente
Carga alta 100MB	31,32	Excelente		28,85	Excelente

Realizado por: Mena Duval, 2021

3.4.3. Comparación del Jitter

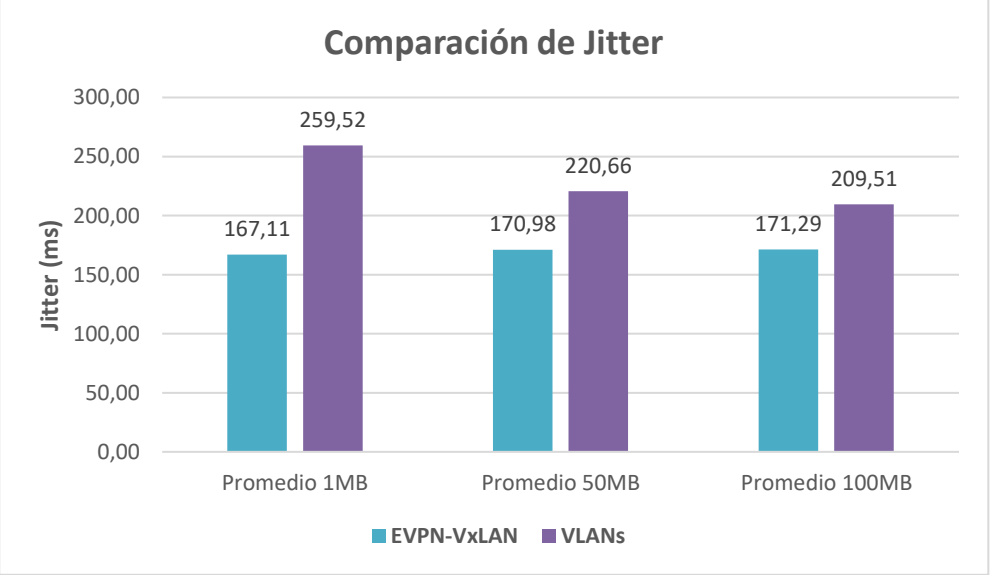

Gráfico 3-3: Comparación entre valores de Jitter promedios de los escenarios propuestos

Realizado por: Mena Duval, 2021

En el grafico 3-3, se muestra los valores de jitter promedio obtenidos en ambos escenarios, en este caso la red con EVPN/VxLAN presenta notablemente mejores parametros de rendimiento en cuanto al jitter en comparacion con los valores obtenidos en la red con VLANs, el motivo de esta

brecha viene ligada con los valores obtenidos en el ancho de banda, ya que normalmente una solucion para mejorar el jitter en las redes es aumentando el ancho de banda de la red, por este motivo gracias al uso de jumbo frames el ancho de banda en la red con EVPN/VxLAN aumenta lo que mejora los valores de jitter de la red.

3.4.4. Comparación de perdida de paquetes

En el grafico 3-4, se observa que en el escenario con EVPN/VxLAN se obtiene una pérdida de paquetes menor en comparación al escenario con VLANs, de igual forma esto se atribuye al uso de tramas gigantes ya que reduce el número de paquetes que se envían a través de la red combinado con un ancho de banda mayor permitiendo aumentar la eficiencia de la red.

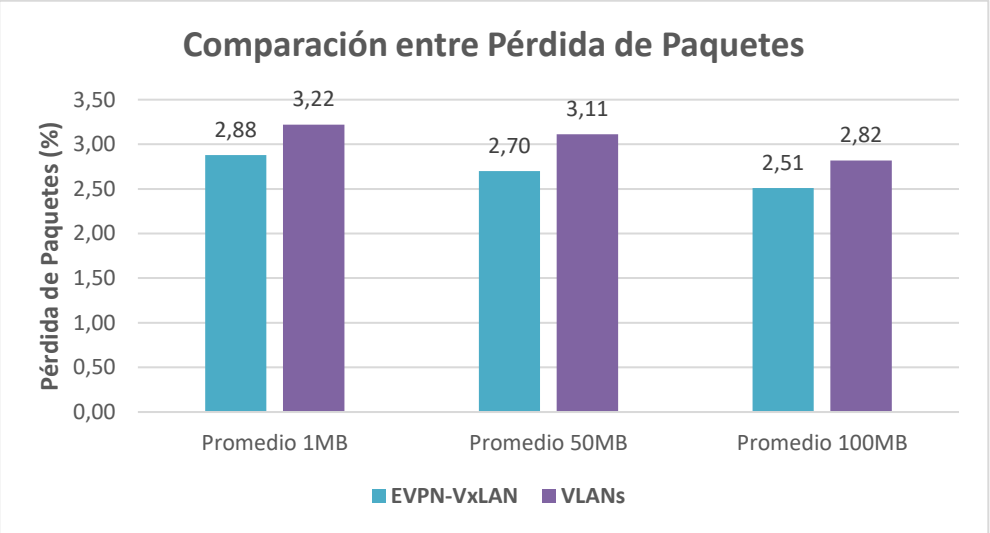

Gráfico 4-3: Comparación entre valores porcentuales de pérdida de paquetes promedios de los escenarios propuestos

Realizado por: Mena Duval, 2021

A continuación, en la tabla 14-3, se realiza la categorización de los valores porcentuales de perdida de paquetes de ambos escenarios.

Tabla 14-3: Categorización de los valores de pérdida de paquetes

Categorización de los valores porcentuales de pérdida de paquetes					
	EVPN-VxLAN (%)			**VLANs (%)**	
Promedio 1MB	2,88	Adecuado		3,22	Conforme
Promedio 50MB	2,70	Adecuado		3,11	Conforme
Promedio 100MB	2,51	Adecuado		2,82	Adecuado

Realizado por: Mena Duval, 2021

CONCLUSIONES

- En base a la investigación realizada acerca de las arquitecturas de centros de datos, se estableció que la arquitectura Clos presenta múltiples mejoras en comparación de la arquitectura tradicional, gracias a que protocolos y mecanismos como bridging y STP ya no son factores principales para la operación, de este modo la arquitectura spine-leaf permite aumentar el ancho de banda por medio de la adición de espinas, aprovechar cada uno de los enlaces gracias a la adopción del protocolo ECMP y adentrarse al mundo virtualizado permitiendo la convivencia entre máquinas virtuales olvidando la preocupación de la sobresuscripción gracias a las propiedades que presenta la tecnología VxLAN.
- Por medio de la investigación acerca de la interconexión de centros de datos mediante BGP EVPN/VxLAN, se concluye que es mejor optar por el uso de un conjunto de hojas de borde como mecanismo de interconexión, ya que estas no requieren planificación y configuraciones adicionales con respecto a las espinas de borde, puesto que al desempeñarse como VTEPs aíslan el interior del centro de datos del mundo exterior logrando así una medida de estabilidad y seguridad; y en caso de que una hoja pierda la conexión con el mundo exterior se produce un impacto menor para el centro de datos en comparación con la pérdida de una espina de borde.
- Se pudo constatar que el uso del plano de control por medio de BGP EVPN evita el uso del mecanismo de inundación y aprendizaje para el descubrimiento de los VTEPs remotos y el anuncio de direcciones MAC, es así que la configuración del plano de control se volvió mucho más simple y presenta un enfoque directo a la automatización del centro de datos.
- En los resultados obtenidos en el parámetro ancho de banda es importante destacar que gracias al uso de jumbo frames en el escenario con EVPN/VxLAN se logra un ancho de banda promedio de 755 Kbps mientras que en el escenario con VLANs implementado con la MTU estándar se obtiene un promedio de 106 Kbps, por lo que se concluye que con el uso de jumbo frames se puede asegurar un mayor rendimiento cumpliendo así uno de los requisitos que exigen los centros de datos.
- Al analizar los parámetros de latencia se observa que la sobrecarga de 50 bytes que genera VxLAN al encapsular los paquetes, incrementa los valores de latencia, debido al consumo de recursos y mayor procesamiento de paquetes; sin embargo al establecer la categorización de estos valores en función de las recomendaciones de los estándares reguladores, se concluye que los parámetros de latencias para la topología EVPN/VxLAN se encuentran categorizados como “Excelentes” por lo que no representa un problema dentro del rendimiento de un centro de datos.

RECOMENDACIONES

- Se recomienda utilizar dentro de las configuraciones el tipo de ruta 5 de EVPN para la salida hacia el mundo exterior mediante L3VNI ya que se basa en una estructura IP-VRF que es propia para la interoperabilidad en capa 3, adicional a esto se recomienda usar el mecanismo IRB simétrico ya que es considerado el más escalable y que trabaja conjuntamente con L3VNI para el enrutamiento y puenteo.
- Analizar detenidamente los requisitos de implementación de VxLAN referente a la sobrecarga que genera el encabezado y las posibles demandas y necesidades que trae consigo.
- Al configurar tramas gigantes en una red se recomienda realizar un análisis de todas las interfaces que necesitan habilitar los marcos gigantes, ya que debido a una mala configuración la red experimentara una degradación parcial o total del rendimiento y a la vez una comunicación inestable o inexistente.
- Se recomienda tomar en cuenta los requisitos de cada uno de los equipos con los que se va a trabajar dentro de la plataforma Gns3 de modo que el consumo del CPU y memoria RAM queden sobrando ya que los propios creadores aseveran que un equipo con mayores recursos permitirá el correcto funcionamiento de las topologías emuladas, caso contrario las limitaciones se verán reflejadas negativamente dentro del desempeño de la red y obtención de resultados.
- Se debe considerar las limitaciones de los equipos dentro de la plataforma Gns3 en cuanto al ancho de banda, si bien es cierto que en promedio se puede alcanzar velocidades de $1 \sim 2\ Mbps$, al configurar los parámetros de los enlaces existe una degradación aun mayor ya que la plataforma tiene que emular cada una de las propiedades de los enlaces lo que genera un mayor consumo de recursos, por ende, el rendimiento en cuanto al ancho de banda se ve afectado notablemente.
- Se recomienda considerar el uso de equipos físicos ya que en vista de las limitaciones que presentan las plataformas de emulación, existen dificultades al momento de la puesta en marcha de las topologías que demandan de muchos recursos, sumado a esto es necesario el uso de tarjetas de red (NIC) de próxima generación que estén familiarizadas con el Overhead de VxLAN y con ello eliminar la pequeña diferencia que existe dentro de los valores de latencias entre redes nativas y superpuestas.

BIBLIOGRAFÍA

ADUSEI, R. Measuring Network Performance With Different Levels Of Firewall Security [en línea] (Master's). Kwame Nkrumah University of Science and Technology. 2016. pp. 64-65. Disponible en: http://ir.knust.edu.gh/bitstream/123456789/10134/1/RICHMOND ADUSEI.pdf.

AL-SHAWI, M. *CCDE Study Guide* [en línea]. Indianapolis-USA: Cisco Press, 2015. ISBN 978-1-58714-461-5. Disponible en: https://www.ciscopress.com/store/ccde-study-guide-9781587144615.

ARIGANELLO, E.; & BARRIENTOS, E. *REDES CISCO CCNP a Fondo* [en línea]. Alfaomega Grupo Editor, 2010. ISBN 978-607-7854-79-1. p. 794. Disponible en: https://www.academia.edu/40840723/REDES_CISCO_CCNP_a_Fondo_Guía_de_estudio_para_profesionales.

ARISTA NETWORKS. *Arista Design Guide Data Center Interconnection with VXLAN* [en línea], 2014. [Consulta: 29 octubre 2021]. Disponible en: https://www.arista.com/assets/data/pdf/Whitepapers/Arista_Design_Guide_DCI_with_VXLAN.pdf.

ARISTA NETWORKS. *EOS 4.27.0F User Manual* [en línea], 2021. [Consulta: 22 octubre 2021]. Disponible en: https://www.arista.com/en/um-eos/eos-evpn-overview.

BATES; et al. *rfc4760* [en línea], 2007. [Consulta: 12 octubre 2021]. Disponible en: https://datatracker.ietf.org/doc/html/rfc4760.

BOOKHAM, C. *Versatile Routing and Services with BGP* [en línea]. Indianapolis-USA: Wiley, 2014. ISBN 978-1-118-87562-9. p. 263. Disponible en: https://www.wiley.com/en-us/Versatile+Routing+and+Services+with+BGP%3A+Understanding+and+Implementing+BGP+in+SR+OS-p-9781118875629.

BUÑAY, P. Aplicación de la Arquitectura DIFFSERV sobre redes MPLS para la provisión de QoS punto a punto en la transmisión de tráfico en tiempo real (Trabajo de titulación). (Maestría): Escuela Superior Politécnica de Chimborazo, Ecuador. 2013. pp. 74-83. Disponible en: http://dspace.espoch.edu.ec/bitstream/123456789/4034/1/20T00464.pdf.

BURESH, B.; et al. *A Modern, Open and Scalable Fabric VXLAN EVPN* [en línea]. San Jose-California: Cisco, 2017. p. 53. [Consulta: 22 octubre 2021]. Disponible en: https://www.cisco.com/c/dam/en/us/td/docs/switches/datacenter/nexus9000/sw/vxlan_evpn/VXLAN_EVPN.pdf.

CAPPUCIO, D. *The Data Center is Dead* [blog]. 2018. [Consulta: 22 diciembre 2020].

Disponible en: https://blogs.gartner.com/david_cappuccio/2018/07/26/the-data-center-is-dead/.

CHACHA, P. Evaluación de una red mpls usando diffserv para mejorar el rendimiento en aplicaciones con QoS [en línea] (Trabajo de titulación). (Maestría) Escuela Superior Politécnica de Chimborazo, Ecuador. 2019. p. 19. [Consulta: 15 Noviembre 2021]. Disponible en: http://dspace.espoch.edu.ec/bitstream/123456789/10833/1/20T01202.pdf.

CHANDRA, D. *This Week: Data Center Deployment With Evpn/Vxlan* [en línea]. USA: Juniper Networks, 2017. ISBN 978-1-941441-58-9, pp. 15-238. Disponible en: https://www.juniper.net/documentation/en_US/day-one-books/TW_DCDeployment.v2.pdf.

CISCO. *IP Routing: BGP Configuration Guide* [en línea]. 2019a. [Consulta: 17 noviembre 2021]. Disponible en: https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute_bgp/configuration/xe-16/irg-xe-16-book/irg-next-hop.html.

CISCO. *What Is a Data Center* [en línea]. 2019b. [Consulta: 10 septiembre 2021]. Disponible en: https://www.cisco.com/c/en/us/solutions/data-center-virtualization/what-is-a-data-center.html.

COMCAST. *Ethernet Network Service Technical Description* [en línea]. 2017. [Consulta: 15 noviembre 2021]. Disponible en: https://business.comcast.com/~/media/business_comcast_com/PDFs/Ethernet Network Services/Ethernet Network Service Technical Description_SLS56950_5.17_REV3.pdf.

COMPUTERWORLD. *Importancia y beneficios de una solución DCI* [en línea]. 2020. [Consulta: 7 octubre 2020]. Disponible en: https://www.computerworld.es/tecnologia/importancia-y-beneficios-de-una-solucion-dci.

DATTA. *Importancia y beneficios de una solución DCI* [en línea]. 2020. [Consulta: 7 octubre 2020]. Disponible en: https://datta.com.ec/articulo/importancia-y-beneficios-de-una-solucion-dci.

DONOHUE, D.; & STEWART, B. *CCNP Routing and Switching Quick Reference* [en línea]. Indianapolis-USA: Cisco Press, 2010. ISBN 978-1-58720-284-1. pp. 69-77. Disponible en: https://www.ciscopress.com/store/ccnp-routing-and-switching-quick-reference-642-902-9780132900065.

DOYLE, J. *Routing TCP/IP CCIE Professional Development* [en línea]. 2ª ed. Indianapolis-USA: Cisco Press, 2016. ISBN 978-1-58705-470-9. pp. 103-618. Disponible en: https://www.ciscopress.com/store/routing-tcp-ip-volume-ii-ccie-professional-development-9781587054709?w_ptgrevartcl=Introduction+to+BGP_2738462.

DUTT, D. *BGP in the Data Center* [en línea]. California-USA.: O'Reilly Media, Inc., 2017. Disponible en: https://learning.oreilly.com/library/view/bgp-in-the/9781491983416/.

DUTT, D. *EVPN in the Data Center* [en línea]. California-USA: O'Reilly Media, Inc., 2018. Disponible en: https://learning.oreilly.com/library/view/evpn-in-the/9781492029045/.

DUTT, D. *Cloud Native Data Center Networking: Architecture, Protocols, and Tools* [en línea]. First Edit. United States of America: O'Reilly Media, Inc., 2019. ISBN 1492045608,978-1492045601. pp. 5-345. Disponible en: https://www.oreilly.com/library/view/cloud-native-data/9781492045595/.

ENTERASYS NETWORKS. "Enterasys Design Center Networking – Connectivity and Topology Design Guide". *Enterasys Networks* [en línea], 2013, p. 8. [Consulta: 17 septiembre 2021]. Disponible en: https://www.computer-pdf.com/network/130-tutorial-data-center-network-design-course.html.

EXTREME NETWORKS. *Extreme IP Fabric Architecture* [en línea], 2018, p. 16. Disponible en: https://kapost-files-prod.s3.amazonaws.com/kapost/55ba7c9e07003d9aab000394/studio/content/5a67611cb0f62c00a10000d4/revisions/1532031743-a9fd7070-48a4-4544-960d-77a474efa6eb/Extreme_IP_Fabric_Architecture-V3_20180519_rebranded.pdf.

FOSCO CONNECT. *WHAT ARE DATA CENTERS?* [en línea]. 2012. [Consulta: 21 noviembre 2021]. Disponible en: https://www.fiberoptics4sale.com/blogs/archive-posts/95041990-what-are-data-centers.

GAI, S. *Building a Future-Proof Cloud Infrastructure* [en línea]. USA: Addison-Wesley Professional, 2020. ISBN 978-0-13-662409-7. pp. 15-22. Disponible en: https://learning.oreilly.com/library/view/building-a-future-proof/9780136624226/.

GLINKOWSKI, M. "Centros De Datos". *ABB review* [en línea], 2013, vol. 4, p. 7. [Consulta: 10 septiembre 2021]. Disponible en: https://library.e.abb.com/public/a53d14ca3eea7516c1257c4000329fe8/Revista ABB 4-2013_72dpi.pdf.

GNS3. *Terrible internet throughput* [en línea]. 2019. Disponible en: https://www.gns3.com/community/featured/terrible-internet-throughput.

GNS3. *Getting Started with GNS3* [en línea]. 2021. [Consulta: 8 noviembre 2021]. Disponible en: https://docs.gns3.com/docs/.

GÓMEZ, M. *Maximizing Performance in VXLAN Overlay Networks* [en línea]. 2017. [Consulta: 18 enero 2022]. Disponible en: https://engineering.telefonica.com/maximizing-

performance-in-vxlan-overlay-networks-ec35ebe29440.

GYARMATHY, K. *Vxchnge* [blog]. 2019. [Consulta: 10 septiembre 2021]. Disponible en: https://www.vxchnge.com/blog/data-center-definition.

HAFIZ, A.; & SUSIANTO, D. *Analysis of Internet Service Quality Using Internet Control Message Protocol* [en línea], 2019. p. 3. Disponible en: https://www.researchgate.net/publication/336805138_Analysis_of_Internet_Service_Quality_Using_Internet_Control_Message_Protocol.

HUAWEI. *BGP Message Format* [en línea]. 2020. [Consulta: 9 octubre 2021]. Disponible en: https://support.huawei.com/enterprise/en/doc/EDOC1100143231/1fd59166/bgp-message-format.

HUAWEI. *Using VXLAN to Implement DCI* [en línea]. 2021. [Consulta: 6 enero 2022]. Disponible en: https://support.huawei.com/enterprise/en/doc/EDOC1100198428/ba47102/using-vxlan-to-implement-dci.

ITU. *Quality of Service REGULATION MANUAL* [en línea]. 2017. ISBN 978-92-61-25791-0. Disponible en: https://www.itu.int/dms_pub/itu-d/opb/pref/D-PREF-BB.QOS_REG01-2017-PDF-E.pdf.

JAIN, V.; & EDGEWORTH, B. *Troubleshooting BGP: A Practical Guide to Understanding and Troubleshooting BGP* [en línea]. Indianapolis-USA: Cisco Press, 2016. ISBN 978-1-58714-464-6. pp. 1-657. Disponible en: https://www.ciscopress.com/store/troubleshooting-bgp-a-practical-guide-to-understanding-9781587144646.

JESUS, J.; et al. MPLS networks for inter substation communication for current differential protection applications in digital substations [en línea]. University of Strathclyde, Glasgow, Reino Unido. 2015. Disponible en: https://strathprints.strath.ac.uk/48807/1/PP021.pdf.

JUNIPER. VXLAN Data Center Interconnect Using EVPN Overview [en línea]. 2021. [Consulta: 6 enero 2022]. Disponible en: https://www.juniper.net/documentation/us/en/software/junos/evpn-vxlan/topics/concept/vxlan-evpn-integration-overview.html.

KAPADIA, S.; et al. *Using TRILL, FabricPath, and VXLAN* [en línea]. Indianapolis-USA: Cisco Press, 2014. ISBN 978-1-58714-393-9. pp. 6-8. Disponible en: https://www.ciscopress.com/store/using-trill-fabricpath-and-vxlan-designing-massively-9780133393040.

KHATIMI, H.; et al. Performance Comparison Between Copper Cables and Fiber Optic in Data

Transfer on Banjarmasin Weather Temperature Conditions. [en línea]. Lambung Mangkurat University, Indonesia. 2019. p. 8. Disponible en: https://www.matec-conferences.org/articles/matecconf/pdf/2019/29/matecconf_icsbe2019_05022.pdf.

KIO NETWORKS. *La correcta gestión de un Data center* [en línea]. [Consulta: 10 septiembre 2021]. ISSN 13697021. p. 2. Disponible en: https://www.kionetworks.com/hubfs/Campaña Inbound Data center/AW/Entregable Awareness.pdf.

KIRAN, T. *Design and Implementation of Transparent Anonymous FTP for Linux* [en línea]. Indian Institute of Technology, Kanpur, 1998. pp. 39-40. Disponible en: https://cseweb.ucsd.edu/~ktati/Papers/MTech_Thesis.pdf.

KRATTIGER, L.; et al. *Building Data Centers with VXLAN BGP EVPN* [en línea]. Indianapolis-USA: Cisco Press, 2017. ISBN 978-1-58714-467-7. Disponible en: https://learning.oreilly.com/library/view/building-data-centers/9780134514895/.

LEWIS, C.; et al. *Selecting MPLS VPN Services* [en línea]. Indianapolis-USA: Cisco Press, 2006. ISBN 978-1-58705-191-3. p. 171. Disponible en: https://www.ciscopress.com/store/selecting-mpls-vpn-services-9781587051913?w_ptgrevartcl=Implementing+Quality+of+Service+Over+Cisco+MPLS+VPNs_471096.

MAHALINGAM, M.; et al. *RFC7348* [en línea]. 2014. [Consulta: 17 octubre 2021]. Disponible en: https://datatracker.ietf.org/doc/html/rfc7348.

MAKOWSKI, L.; et al. Evaluation of virtualization and traffic filtering methods for container networks [en línea]. University of Amsterdam, Amsterdam, Países Bajos. 2019 : Disponible en: https://scinet.supercomputing.org/community/documents/46/sc17-Makowski-container_networking.pdf.

MARSCHKE, D.; et al. *Junos Enterprise Routing* [en línea]. 2ª ed. United States of America: O'Reilly Media, Inc., 2011. ISBN 978-1-449-39863-7. pp. 33-37. Disponible en: https://learning.oreilly.com/library/view/junos-enterprise-routing/9781449309633/.

NAUFAL, Y. *Performance Analysis Of Vxlan-Evpn Protocol Using Open Network On Overlay Network* [en línea]. Institut Teknologi Telkom Purwokerto. 2021. Disponible en: http://repository.ittelkom-pwt.ac.id/6489/.

NELSON, W. "Introduction to Spine-Leaf Networking Designs". *Lenovo Press* [en línea], (2017), (United States of America), p. 6. [Consulta: 25 septiembre 2021]. Disponible en: https://lenovopress.com/lp0573.pdf.

NOGHANI, K.; & KASSLER, A. SDN Enhanced Ethernet VPN for Data Center Interconnect.

[en línea]. Karlstad University. 2019. Disponible en: https://arxiv.org/pdf/1911.00779.pdf.

NORTHLAND COMMUNICATIONS. *Northland Communications Dedicated Internet, Cloud and MPLS Services* [en línea]. 2013. [Consulta: 15 noviembre 2021]. Disponible en: https://docplayer.net/3018486-Northland-communications-dedicated-internet-cloud-and-mpls-services-service-level-agreement-sla-02-07-2013.html.

NOURI, A. *Throughput analysis in GNS3* [en línea]. 2017. [Consulta: 22 noviembre 2021]. Disponible en: https://gns3.com/community/blog/bandwidth-analysis-in-gns3.

PASANEN, T. *Virtual Extensible LAN (VXLAN) A Practical guide to VXLAN solution* [en línea]. 2019. [Consulta: 7 noviembre 2021]. Disponible en: https://leanpub.com/virtualextensiblelanvxlanapracticalguidetovxlansolutionpart1.

PEREZ, A. *IP, Ethernet and MPLS Networks: Resource and Fault Management* [en línea]. USA: Wiley, 2011. ISBN 978-1-84821-285-5. Disponible en: http://www.iste.co.uk/book.php?id=393.

PLURIBUS NETWORKS; & DELL EMC. *Next-Generation Data Center Interconnect Powered by the Pluribus Adaptive Cloud Fabric* [en línea]. 2018. [Consulta: 29 octubre 2021]. Disponible en: https://www.pluribusnetworks.com/assets/Dell-Pluribus-NextGen-DCI-SO-042518.pdf.

REKHTER; et al. *rfc4271* [en línea]. 2006. [Consulta: 8 octubre 2021]. Disponible en: https://datatracker.ietf.org/doc/html/rfc4271#section-5.1.6.

SHAMSEE, N.; et al. *CCNA Data Center DCICT 640-916 Official Cert Guide* [en línea]. Indianapolis-USA: Cisco Press, 2015. ISBN 978-1-58714-422-6. Disponible en: https://www.ciscopress.com/store/ccna-data-center-dcict-640-916-official-cert-guide-9781587144226.

SINGH, T.; et al. *VXLAN and EVPN for data center network transformation* [en línea]. Delhi, 2017. Disponible en: https://ieeexplore.ieee.org/document/8203947.

SULLIVAN, K. Top Five Attributes Of A Successful Data Center [en línea]. 2016. [Consulta: 13 septiembre 2021]. Disponible en: https://www.forbes.com/sites/centurylink/2016/05/11/top-five-attributes-of-a-successful-data-center/?sh=2d2c039a3f23.

TAKAMÄKI, M., 2018. Overlay Technologies and Microsegmentation in Data Centers [en línea]. School of Technology, Communication and Transport, 2018. Disponible en: https://www.theseus.fi/bitstream/handle/10024/146657/Takamaki_Masi_Final.pdf?sequence=1.

VARNUM, D. *Arista BGP EVPN – Overview and Concepts* [en línea]. 2018. [Consulta: 30 octubre 2021]. Disponible en: https://overlaid.net/2018/08/27/arista-bgp-evpn-overview-and-concepts/.

VMWARE. *Guía de instalación de Cross-vCenter NSX* [en línea]. 2021. [Consulta: 19 octubre 2021]. Disponible en: https://docs.vmware.com/es/VMware-NSX-Data-Center-for-vSphere/6.4/nsx_64_cross_vc_install.pdf.

WRIGHTSON, C. *Day One: Data Center Fundamentals* [en línea]. USA: Juniper Networks, 2016. ISBN 978-1-941441-40-4. Disponible en: https://www.juniper.net/documentation/en_US/day-one-books/DC_Fundamentals.pdf.

ZHANG, R.; & BARTELL, M. *BGP Design and Implementation* [en línea]. Indianapolis-USA: Cisco Press, 2016. ISBN 1-58705-109-5. Disponible en: https://www.ciscopress.com/store/bgp-design-and-implementation-9780133433555.

espoch | Dirección de Bibliotecas y Recursos del Aprendizaje

UNIDAD DE PROCESOS TÉCNICOS Y ANÁLISIS BIBLIOGRÁFICO Y DOCUMENTAL

REVISIÓN DE NORMAS TÉCNICAS, RESUMEN Y BIBLIOGRAFÍA

Fecha de entrega: 12 / 04 / 2022

INFORMACIÓN DEL AUTOR/A (S)
Nombres – Apellidos: DUVAL ANDRÉS MENA PAREDES
INFORMACIÓN INSTITUCIONAL
Facultad: INFORMÁTICA Y ELECTRÓNICA
Carrera: TELECOMUNICACIONES
Título a optar: INGENIERO EN ELECTRÓNICA, TELECOMUNICACIONES Y REDES
f. Analista de Biblioteca responsable: Lcdo. Holger Ramos, MSc.

UNIDAD DE PROCESOS TÉCNICOS DBRA DIRECCIÓN DE BIBLIOTECAS Y RECURSOS DEL APRENDIZAJE ESPOCH

Firmado electrónicamente por: HOLGER GERMAN RAMOS UVIDIA

0549-DBRA-UPT-2022